



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2013

---

## **Search for the production of ZW and ZZ boson pairs decaying into charged leptons and jets in proton-antiproton collisions at $\sqrt{s}=1.96$ TeV**

CDF Collaboration ; et al ; Canelli, F ; Kilminster, B

**Abstract:** We present a measurement of the production cross section for ZW and ZZ boson pairs in final states with a pair of charged leptons, from the decay of a Z boson, and at least two jets, from the decay of a W or Z boson, using the full sample of proton-antiproton collisions recorded with the CDF II detector at the Tevatron, corresponding to  $8.9 \text{ fb}^{-1}$  of integrated luminosity. We increase the sensitivity to vector-boson decays into pairs of quarks using a neural-network discriminant that exploits the differences between the spatial spread of energy depositions and charged-particle momenta contained within the jet of particles originating from quarks and gluons. Additionally, we employ new jet energy corrections to Monte Carlo simulations that account for differences in the observed energy scales for quark and gluon jets. The number of signal events is extracted through a simultaneous fit to the dijet mass spectrum in three classes of events: events likely to contain jets with a heavy-quark decay, events likely to contain jets originating from light quarks, and events that fail these identification criteria. We determine the production cross section to be  $\text{ZW}+\text{ZZ}=2.5+2.0-1.0 \text{ pb}$  ( $<6.1 \text{ pb}$  at the 95% confidence level), consistent with the standard model prediction of  $5.1 \text{ pb}$ .

DOI: <https://doi.org/10.1103/PhysRevD.88.092002>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-92069>

Journal Article

Accepted Version

Originally published at:

CDF Collaboration; et al; Canelli, F; Kilminster, B (2013). Search for the production of ZW and ZZ boson pairs decaying into charged leptons and jets in proton-antiproton collisions at  $\sqrt{s}=1.96$  TeV. Physical Review D (Particles, Fields, Gravitation and Cosmology), 88(9):092002.

DOI: <https://doi.org/10.1103/PhysRevD.88.092002>

# Search for the production of $ZW$ and $ZZ$ boson pairs decaying into charged leptons and jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV

T. Aaltonen,<sup>21</sup> S. Amerio<sup>jj,39</sup> D. Amidei,<sup>31</sup> A. Anastassov<sup>v,15</sup> A. Annovi,<sup>17</sup> J. Antos,<sup>12</sup> G. Apollinari,<sup>15</sup> J.A. Appel,<sup>15</sup> T. Arisawa,<sup>52</sup> A. Artikov,<sup>13</sup> J. Asaadi,<sup>47</sup> W. Ashmanskas,<sup>15</sup> B. Auerbach,<sup>2</sup> A. Aurisano,<sup>47</sup> F. Azfar,<sup>38</sup> W. Badgett,<sup>15</sup> T. Bae,<sup>25</sup> A. Barbaro-Galtieri,<sup>26</sup> V.E. Barnes,<sup>43</sup> B.A. Barnett,<sup>23</sup> P. Barria<sup>ll,41</sup> P. Bartos,<sup>12</sup> M. Baucé<sup>jj,39</sup> F. Bedeschi,<sup>41</sup> S. Behari,<sup>15</sup> G. Bellettini<sup>kk,41</sup> J. Bellinger,<sup>54</sup> D. Benjamin,<sup>14</sup> A. Beretvas,<sup>15</sup> A. Bhatti,<sup>45</sup> K.R. Bland,<sup>5</sup> B. Blumenfeld,<sup>23</sup> A. Bocci,<sup>14</sup> A. Bodek,<sup>44</sup> D. Bortoletto,<sup>43</sup> J. Boudreau,<sup>42</sup> A. Boveia,<sup>11</sup> L. Brigliadori<sup>ii,6</sup> C. Bromberg,<sup>32</sup> E. Brucken,<sup>21</sup> J. Budagov,<sup>13</sup> H.S. Budd,<sup>44</sup> K. Burkett,<sup>15</sup> G. Busetto<sup>jj,39</sup> P. Bussey,<sup>19</sup> P. Butti<sup>kk,41</sup> A. Buzatu,<sup>19</sup> A. Calamba,<sup>10</sup> S. Camarda,<sup>4</sup> M. Campanelli,<sup>28</sup> F. Canelli<sup>cc,11</sup> B. Carls,<sup>22</sup> D. Carlsmith,<sup>54</sup> R. Carosi,<sup>41</sup> S. Carrillo<sup>l,16</sup> B. Casal<sup>j,9</sup> M. Casarsa,<sup>48</sup> A. Castro<sup>ii,6</sup> P. Catastini,<sup>20</sup> D. Cauz<sup>qrrr,48</sup> V. Cavaliere,<sup>22</sup> M. Cavalli-Sforza,<sup>4</sup> A. Cerri<sup>e,26</sup> L. Cerrito<sup>q,28</sup> Y.C. Chen,<sup>1</sup> M. Chertok,<sup>7</sup> G. Chiarelli,<sup>41</sup> G. Chlachidze,<sup>15</sup> K. Cho,<sup>25</sup> D. Chokheli,<sup>13</sup> A. Clark,<sup>18</sup> C. Clarke,<sup>53</sup> M.E. Convery,<sup>15</sup> J. Conway,<sup>7</sup> M. Corbo<sup>y,15</sup> M. Cordelli,<sup>17</sup> C.A. Cox,<sup>7</sup> D.J. Cox,<sup>7</sup> M. Cremonesi,<sup>41</sup> D. Cruz,<sup>47</sup> J. Cuevas<sup>x,9</sup> R. Culbertson,<sup>15</sup> N. d'Ascenzo<sup>u,15</sup> M. Datta<sup>ff,15</sup> P. de Barbaro,<sup>44</sup> L. Demortier,<sup>45</sup> M. Deninno,<sup>6</sup> M. D'Errico<sup>jj,39</sup> F. Devoto,<sup>21</sup> A. Di Canto<sup>kk,41</sup> B. Di Ruzza<sup>p,15</sup> J.R. Dittmann,<sup>5</sup> S. Donati<sup>kk,41</sup> M. D'Onofrio,<sup>27</sup> M. Dorigo<sup>ss,48</sup> A. Driutti<sup>qrrr,48</sup> K. Ebina,<sup>52</sup> R. Edgar,<sup>31</sup> A. Elagin,<sup>47</sup> R. Erbacher,<sup>7</sup> S. Errede,<sup>22</sup> B. Esham,<sup>22</sup> S. Farrington,<sup>38</sup> J.P. Fernández Ramos,<sup>29</sup> R. Field,<sup>16</sup> G. Flanagan<sup>s,15</sup> R. Forrest,<sup>7</sup> M. Franklin,<sup>20</sup> J.C. Freeman,<sup>15</sup> H. Frisch,<sup>11</sup> Y. Funakoshi,<sup>52</sup> C. Galloni<sup>kk,41</sup> A.F. Garfinkel,<sup>43</sup> P. Garosi<sup>ll,41</sup> H. Gerberich,<sup>22</sup> E. Gerchtein,<sup>15</sup> S. Giagu,<sup>46</sup> V. Giakoumopoulou,<sup>3</sup> K. Gibson,<sup>42</sup> C.M. Ginsburg,<sup>15</sup> N. Giokaris,<sup>3</sup> P. Giromini,<sup>17</sup> G. Giurgiu,<sup>23</sup> V. Glagolev,<sup>13</sup> D. Glenzinski,<sup>15</sup> M. Gold,<sup>34</sup> D. Goldin,<sup>47</sup> A. Golossanov,<sup>15</sup> G. Gomez,<sup>9</sup> G. Gomez-Ceballos,<sup>30</sup> M. Goncharov,<sup>30</sup> O. González López,<sup>29</sup> I. Gorelov,<sup>34</sup> A.T. Goshaw,<sup>14</sup> K. Goulianos,<sup>45</sup> E. Gramellini,<sup>6</sup> S. Grinstein,<sup>4</sup> C. Grosso-Pilcher,<sup>11</sup> R.C. Group,<sup>51,15</sup> J. Guimaraes da Costa,<sup>20</sup> S.R. Hahn,<sup>15</sup> J.Y. Han,<sup>44</sup> F. Happacher,<sup>17</sup> K. Hara,<sup>49</sup> M. Hare,<sup>50</sup> R.F. Harr,<sup>53</sup> T. Harrington-Taber<sup>m,15</sup> K. Hatakeyama,<sup>5</sup> C. Hays,<sup>38</sup> J. Heinrich,<sup>40</sup> M. Herndon,<sup>54</sup> A. Hocker,<sup>15</sup> Z. Hong,<sup>47</sup> W. Hopkins<sup>f,15</sup> S. Hou,<sup>1</sup> R.E. Hughes,<sup>35</sup> U. Husemann,<sup>55</sup> M. Hussein<sup>aa,32</sup> J. Huston,<sup>32</sup> G. Introzzi<sup>nnoo,41</sup> M. Iori<sup>pp,46</sup> A. Ivanov<sup>o,7</sup> E. James,<sup>15</sup> D. Jang,<sup>10</sup> B. Jayatilaka,<sup>15</sup> E.J. Jeon,<sup>25</sup> S. Jindariani,<sup>15</sup> M. Jones,<sup>43</sup> K.K. Joo,<sup>25</sup> S.Y. Jun,<sup>10</sup> T.R. Junk,<sup>15</sup> M. Kambeitz,<sup>24</sup> T. Kamon,<sup>25,47</sup> P.E. Karchin,<sup>53</sup> A. Kasmi,<sup>5</sup> Y. Kato<sup>n,37</sup> W. Ketchum<sup>gg,11</sup> J. Keung,<sup>40</sup> B. Kilminster<sup>cc,15</sup> D.H. Kim,<sup>25</sup> H.S. Kim,<sup>25</sup> J.E. Kim,<sup>25</sup> M.J. Kim,<sup>17</sup> S.H. Kim,<sup>49</sup> S.B. Kim,<sup>25</sup> Y.J. Kim,<sup>25</sup> Y.K. Kim,<sup>11</sup> N. Kimura,<sup>52</sup> M. Kirby,<sup>15</sup> K. Knoepfel,<sup>15</sup> K. Kondo,<sup>52,\*</sup> D.J. Kong,<sup>25</sup> J. Konigsberg,<sup>16</sup> A.V. Kotwal,<sup>14</sup> M. Kreps,<sup>24</sup> J. Kroll,<sup>40</sup> M. Kruse,<sup>14</sup> T. Kuhr,<sup>24</sup> M. Kurata,<sup>49</sup> A.T. Laasanen,<sup>43</sup> S. Lammel,<sup>15</sup> M. Lancaster,<sup>28</sup> K. Lannon<sup>w,35</sup> G. Latino<sup>ll,41</sup> H.S. Lee,<sup>25</sup> J.S. Lee,<sup>25</sup> S. Leo,<sup>41</sup> S. Leone,<sup>41</sup> J.D. Lewis,<sup>15</sup> A. Limosani<sup>r,14</sup> E. Lipeles,<sup>40</sup> A. Lister<sup>a,18</sup> H. Liu,<sup>51</sup> Q. Liu,<sup>43</sup> T. Liu,<sup>15</sup> S. Lockwitz,<sup>55</sup> A. Loginov,<sup>55</sup> D. Lucchesi<sup>jj,39</sup> A. Lucà,<sup>17</sup> J. Lueck,<sup>24</sup> P. Lujan,<sup>26</sup> P. Lukens,<sup>15</sup> G. Lungu,<sup>45</sup> J. Lys,<sup>26</sup> R. Lysak<sup>d,12</sup> R. Madrak,<sup>15</sup> P. Maestro<sup>ll,41</sup> S. Malik,<sup>45</sup> G. Manca<sup>b,27</sup> A. Manousakis-Katsikakis,<sup>3</sup> L. Marchese<sup>hh,6</sup> F. Margaroli,<sup>46</sup> P. Marino<sup>mm,41</sup> M. Martínez,<sup>4</sup> K. Matera,<sup>22</sup> M.E. Mattson,<sup>53</sup> A. Mazzacane,<sup>15</sup> P. Mazzanti,<sup>6</sup> R. McNulty<sup>i,27</sup> A. Mehta,<sup>27</sup> P. Mehtala,<sup>21</sup> C. Mesropian,<sup>45</sup> T. Miao,<sup>15</sup> D. Mietlicki,<sup>31</sup> A. Mitra,<sup>1</sup> H. Miyake,<sup>49</sup> S. Moed,<sup>15</sup> N. Moggi,<sup>6</sup> C.S. Moon<sup>y,15</sup> R. Moore<sup>ddee,15</sup> M.J. Morello<sup>mm,41</sup> A. Mukherjee,<sup>15</sup> Th. Muller,<sup>24</sup> P. Murat,<sup>15</sup> M. Mussini<sup>ii,6</sup> J. Nachtman<sup>m,15</sup> Y. Nagai,<sup>49</sup> J. Naganoma,<sup>52</sup> I. Nakano,<sup>36</sup> A. Napier,<sup>50</sup> J. Nett,<sup>47</sup> C. Neu,<sup>51</sup> T. Nigmanov,<sup>42</sup> L. Nodulman,<sup>2</sup> S.Y. Noh,<sup>25</sup> O. Norniella,<sup>22</sup> L. Oakes,<sup>38</sup> S.H. Oh,<sup>14</sup> Y.D. Oh,<sup>25</sup> I. Oksuzian,<sup>51</sup> T. Okusawa,<sup>37</sup> R. Orava,<sup>21</sup> L. Ortolan,<sup>4</sup> C. Pagliarone,<sup>48</sup> E. Palencia<sup>e,9</sup> P. Palni,<sup>34</sup> V. Papadimitriou,<sup>15</sup> W. Parker,<sup>54</sup> G. Pauletta<sup>qrrr,48</sup> M. Paulini,<sup>10</sup> C. Paus,<sup>30</sup> T.J. Phillips,<sup>14</sup> G. Piacentino,<sup>41</sup> E. Pianori,<sup>40</sup> J. Pilot,<sup>7</sup> K. Pitts,<sup>22</sup> C. Plager,<sup>8</sup> L. Pondrom,<sup>54</sup> S. Poprocki<sup>f,15</sup> K. Potamianos,<sup>26</sup> A. Pranko,<sup>26</sup> F. Prokoshin<sup>z,13</sup> F. Ptohos<sup>g,17</sup> G. Punzi<sup>kk,41</sup> N. Ranjan,<sup>43</sup> I. Redondo Fernández,<sup>29</sup> P. Renton,<sup>38</sup> M. Rescigno,<sup>46</sup> F. Rimondi,<sup>6,\*</sup> L. Ristori,<sup>41,15</sup> A. Robson,<sup>19</sup> T. Rodriguez,<sup>40</sup> S. Rolli<sup>h,50</sup> M. Ronzani<sup>kk,41</sup> R. Roser,<sup>15</sup> J.L. Rosner,<sup>11</sup> F. Ruffini<sup>ll,41</sup> A. Ruiz,<sup>9</sup> J. Russ,<sup>10</sup> V. Rusu,<sup>15</sup> W.K. Sakumoto,<sup>44</sup> Y. Sakurai,<sup>52</sup> L. Santi<sup>qrrr,48</sup> K. Sato,<sup>49</sup> V. Saveliev<sup>u,15</sup> A. Savoy-Navarro<sup>y,15</sup> P. Schlabach,<sup>15</sup> E.E. Schmidt,<sup>15</sup> T. Schwarz,<sup>31</sup> L. Scodellaro,<sup>9</sup> F. Scuri,<sup>41</sup> S. Seidel,<sup>34</sup> Y. Seiya,<sup>37</sup> A. Semenov,<sup>13</sup> F. Sforza<sup>kk,41</sup> S.Z. Shalhout,<sup>7</sup> T. Shears,<sup>27</sup> P.F. Shepard,<sup>42</sup> M. Shimojima<sup>t,49</sup> M. Shochet,<sup>11</sup> I. Shreyber-Tecker,<sup>33</sup> A. Simonenko,<sup>13</sup> K. Sliwa,<sup>50</sup> J.R. Smith,<sup>7</sup> F.D. Snider,<sup>15</sup> H. Song,<sup>42</sup> V. Sorin,<sup>4</sup> R. St. Denis,<sup>19</sup> M. Stancari,<sup>15</sup> D. Stentz<sup>v,15</sup> J. Strologas,<sup>34</sup> Y. Sudo,<sup>49</sup> A. Sukhanov,<sup>15</sup> I. Suslov,<sup>13</sup> K. Takemasa,<sup>49</sup> Y. Takeuchi,<sup>49</sup> J. Tang,<sup>11</sup> M. Tecchio,<sup>31</sup> P.K. Teng,<sup>1</sup> J. Thom<sup>f,15</sup> E. Thomson,<sup>40</sup> V. Thukral,<sup>47</sup> D. Toback,<sup>47</sup> S. Tokar,<sup>12</sup> K. Tollefson,<sup>32</sup> T. Tomura,<sup>49</sup> D. Tonelli<sup>e,15</sup> S. Torre,<sup>17</sup> D. Torretta,<sup>15</sup> P. Totaro,<sup>39</sup> M. Trovato<sup>mm,41</sup> F. Ukegawa,<sup>49</sup> S. Uozumi,<sup>25</sup> F. Vázquez<sup>l,16</sup> G. Velev,<sup>15</sup> C. Vellidis,<sup>15</sup> C. Vernieri<sup>mm,41</sup> M. Vidal,<sup>43</sup> R. Vilar,<sup>9</sup> J. Vizán<sup>bb,9</sup> M. Vogel,<sup>34</sup> G. Volpi,<sup>17</sup> P. Wagner,<sup>40</sup> R. Wallnj<sup>j,15</sup> S.M. Wang,<sup>1</sup> D. Waters,<sup>28</sup> W.C. Wester III,<sup>15</sup> D. Whiteson<sup>c,40</sup> A.B. Wicklund,<sup>2</sup> S. Wilbur,<sup>7</sup> H.H. Williams,<sup>40</sup> J.S. Wilson,<sup>31</sup> P. Wilson,<sup>15</sup> B.L. Winer,<sup>35</sup> P. Wittich<sup>f,15</sup> S. Wolbers,<sup>15</sup> H. Wolfe,<sup>35</sup> T. Wright,<sup>31</sup> X. Wu,<sup>18</sup> Z. Wu,<sup>5</sup> K. Yamamoto,<sup>37</sup> D. Yamato,<sup>37</sup> T. Yang,<sup>15</sup> U.K. Yang,<sup>25</sup> Y.C. Yang,<sup>25</sup> W.-M. Yao,<sup>26</sup> G.P. Yeh,<sup>15</sup> K. Yi<sup>m,15</sup> J. Yoh,<sup>15</sup> K. Yorita,<sup>52</sup> T. Yoshida<sup>k,37</sup> G.B. Yu,<sup>14</sup> I. Yu,<sup>25</sup> A.M. Zanetti,<sup>48</sup> Y. Zeng,<sup>14</sup> C. Zhou,<sup>14</sup> and S. Zucchelli<sup>ii6</sup>

(CDF Collaboration)<sup>†</sup>

<sup>†</sup>Institute of Physics, Academia Sinica, Taipei, Taiwan 11529, Republic of China

- <sup>2</sup>Argonne National Laboratory, Argonne, Illinois 60439, USA  
<sup>3</sup>University of Athens, 157 71 Athens, Greece  
<sup>4</sup>Institut de Física d'Altes Energies, ICREA, Universitat Autònoma de Barcelona, E-08193, Bellaterra (Barcelona), Spain  
<sup>5</sup>Baylor University, Waco, Texas 76798, USA  
<sup>6</sup>Istituto Nazionale di Fisica Nucleare Bologna, <sup>ii</sup>University of Bologna, I-40127 Bologna, Italy  
<sup>7</sup>University of California, Davis, Davis, California 95616, USA  
<sup>8</sup>University of California, Los Angeles, Los Angeles, California 90024, USA  
<sup>9</sup>Instituto de Física de Cantabria, CSIC-University of Cantabria, 39005 Santander, Spain  
<sup>10</sup>Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA  
<sup>11</sup>Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637, USA  
<sup>12</sup>Comenius University, 842 48 Bratislava, Slovakia; Institute of Experimental Physics, 040 01 Kosice, Slovakia  
<sup>13</sup>Joint Institute for Nuclear Research, RU-141980 Dubna, Russia  
<sup>14</sup>Duke University, Durham, North Carolina 27708, USA  
<sup>15</sup>Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA  
<sup>16</sup>University of Florida, Gainesville, Florida 32611, USA  
<sup>17</sup>Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, I-00044 Frascati, Italy  
<sup>18</sup>University of Geneva, CH-1211 Geneva 4, Switzerland  
<sup>19</sup>Glasgow University, Glasgow G12 8QQ, United Kingdom  
<sup>20</sup>Harvard University, Cambridge, Massachusetts 02138, USA  
<sup>21</sup>Division of High Energy Physics, Department of Physics, University of Helsinki, FIN-00014, Helsinki, Finland; Helsinki Institute of Physics, FIN-00014, Helsinki, Finland  
<sup>22</sup>University of Illinois, Urbana, Illinois 61801, USA  
<sup>23</sup>The Johns Hopkins University, Baltimore, Maryland 21218, USA  
<sup>24</sup>Institut für Experimentelle Kernphysik, Karlsruhe Institute of Technology, D-76131 Karlsruhe, Germany  
<sup>25</sup>Center for High Energy Physics: Kyungpook National University, Daegu 702-701, Korea; Seoul National University, Seoul 151-742, Korea; Sungkyunkwan University, Suwon 440-746, Korea; Korea Institute of Science and Technology Information, Daejeon 305-806, Korea; Chonnam National University, Gwangju 500-757, Korea; Chonbuk National University, Jeonju 561-756, Korea; Ewha Womans University, Seoul, 120-750, Korea  
<sup>26</sup>Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA  
<sup>27</sup>University of Liverpool, Liverpool L69 7ZE, United Kingdom  
<sup>28</sup>University College London, London WC1E 6BT, United Kingdom  
<sup>29</sup>Centro de Investigaciones Energéticas Medioambientales y Tecnológicas, E-28040 Madrid, Spain  
<sup>30</sup>Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA  
<sup>31</sup>University of Michigan, Ann Arbor, Michigan 48109, USA  
<sup>32</sup>Michigan State University, East Lansing, Michigan 48824, USA  
<sup>33</sup>Institution for Theoretical and Experimental Physics, ITEP, Moscow 117259, Russia  
<sup>34</sup>University of New Mexico, Albuquerque, New Mexico 87131, USA  
<sup>35</sup>The Ohio State University, Columbus, Ohio 43210, USA  
<sup>36</sup>Okayama University, Okayama 700-8530, Japan  
<sup>37</sup>Osaka City University, Osaka 558-8585, Japan  
<sup>38</sup>University of Oxford, Oxford OX1 3RH, United Kingdom  
<sup>39</sup>Istituto Nazionale di Fisica Nucleare, Sezione di Padova, <sup>jj</sup>University of Padova, I-35131 Padova, Italy  
<sup>40</sup>University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA  
<sup>41</sup>Istituto Nazionale di Fisica Nucleare Pisa, <sup>kk</sup>University of Pisa, <sup>ll</sup>University of Siena, <sup>mm</sup>Scuola Normale Superiore, I-56127 Pisa, Italy, <sup>nn</sup>INFN Pavia, I-27100 Pavia, Italy, <sup>oo</sup>University of Pavia, I-27100 Pavia, Italy  
<sup>42</sup>University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA  
<sup>43</sup>Purdue University, West Lafayette, Indiana 47907, USA  
<sup>44</sup>University of Rochester, Rochester, New York 14627, USA  
<sup>45</sup>The Rockefeller University, New York, New York 10065, USA  
<sup>46</sup>Istituto Nazionale di Fisica Nucleare, Sezione di Roma I, <sup>pp</sup>Sapienza Università di Roma, I-00185 Roma, Italy  
<sup>47</sup>Mitchell Institute for Fundamental Physics and Astronomy, Texas A&M University, College Station, Texas 77843, USA  
<sup>48</sup>Istituto Nazionale di Fisica Nucleare Trieste, <sup>qq</sup>Gruppo Collegato di Udine, <sup>rr</sup>University of Udine, I-33100 Udine, Italy, <sup>ss</sup>University of Trieste, I-34127 Trieste, Italy  
<sup>49</sup>University of Tsukuba, Tsukuba, Ibaraki 305, Japan  
<sup>50</sup>Tufts University, Medford, Massachusetts 02155, USA  
<sup>51</sup>University of Virginia, Charlottesville, Virginia 22906, USA  
<sup>52</sup>Waseda University, Tokyo 169, Japan  
<sup>53</sup>Wayne State University, Detroit, Michigan 48201, USA  
<sup>54</sup>University of Wisconsin, Madison, Wisconsin 53706, USA  
<sup>55</sup>Yale University, New Haven, Connecticut 06520, USA

(Dated: September 30, 2013)

We present a measurement of the production cross section for  $ZW$  and  $ZZ$  boson pairs in final states with a pair of charged leptons, from the decay of a  $Z$  boson, and at least two jets, from the decay of a  $W$  or  $Z$  boson, using the full sample of proton-antiproton collisions recorded with the CDF II detector at the Tevatron, corresponding to  $8.9 \text{ fb}^{-1}$  of integrated luminosity. We increase the sensitivity to vector boson decays into pairs of quarks using a neural network discriminant that exploits the differences between the spatial spread of energy depositions and charged-particle momenta contained within the jet of particles originating from quarks and gluons. Additionally, we employ new jet energy corrections to Monte Carlo simulations that account for differences in the observed energy scales for quark and gluon jets. The number of signal events is extracted through a simultaneous fit to the dijet mass spectrum in three classes of events: events likely to contain jets with a heavy-quark decay, events likely to contain jets originating from light quarks, and events that fail these identification criteria. We determine the production cross section to be  $\sigma_{ZW+ZZ} = 2.5_{-1.0}^{+2.0} \text{ pb}$  ( $< 6.1 \text{ pb}$  at the 95% confidence level), consistent with the standard model prediction of  $5.1 \text{ pb}$ .

PACS numbers:

## I. INTRODUCTION

The standard model (SM) offers precise predictions of production rates associated with self-interactions of the gauge bosons [1]. Differences between these predictions and measured diboson production cross sections may indicate the presence of non-SM physics [2, 3], even specifically in hadronic final states [4]. Additionally, since hadronic final states in diboson production are similar to those from associated Higgs boson production ( $p\bar{p} \rightarrow VH + X$  where  $V=W, Z$ ), the analysis techniques used to measure diboson production in partially hadronic final states are relevant to searches for associated Higgs boson production.

Measurements of diboson production are typically difficult due to the small production cross sections of the order of 10

pb or less [1]. Furthermore, measurements of decay channels where one  $W$  or  $Z$  boson decays hadronically are particularly challenging at hadron colliders: although expected event yields are larger than those in purely leptonic decay channels due to the higher hadronic decay ( $V \rightarrow qq'$ ) branching ratio, the expected backgrounds from QCD multijet processes and  $V$  + jets production are also much greater. Experiments at the Tevatron have previously measured the cross sections for pair production of gauge bosons in partially hadronic decay channels [5–8], but all of these measurements have included contributions from  $WW$  production, which has a higher cross section than that for combined  $ZW$  and  $ZZ$  production. Searches using identification of  $b$ -quark decays in the final states ( $b$ -tagging) to increase sensitivity to events with  $Z \rightarrow b\bar{b}$  decays have been performed [9], but have not yet provided observations of  $ZV$  production in partially hadronic decay channels.

We present a study of  $ZV$  production from a final state with two leptons and at least two jets [10]. We require the two leptons to originate from the decay of a  $Z$  boson and search for associated  $V \rightarrow qq'$  decays by performing a fit to the dijet invariant mass ( $m_{jj}$ ) spectrum of the two leading- $E_T$  [11] jets. To maximize sensitivity to diboson production, we separate events into three channels: a heavy-flavor-tagged channel, largely sensitive to  $ZZ \rightarrow \ell^+ \ell^- b\bar{b}$  decays; a light-flavor-tagged channel, which uses a new artificial-neural-network-based discriminant to preferentially select events with quark-like jets over gluon-like jets; and an untagged channel, which contains the remaining events that pass the event selection requirements. The final fit to the  $m_{jj}$  spectra is performed simultaneously across all three channels.

This paper is organized as follows: in Sec. II we briefly describe the CDF II detector; in Sec. III we describe the data sets and event selection requirements that are used in the  $ZV$  search; in Sec. IV we show the derivation of new jet energy corrections to Monte Carlo simulations that account for differences in the observed energy scales of quark and gluon jets; in Sec. V we provide details of a new neural network-based discriminant that identifies jets more likely to originate from quarks than from gluons; and, in Sec. VI, we describe the signal-extraction method, and report the results of the  $ZV$  search.

---

\*Deceased

<sup>†</sup>With visitors from <sup>a</sup>University of British Columbia, Vancouver, BC V6T 1Z1, Canada, <sup>b</sup>Istituto Nazionale di Fisica Nucleare, Sezione di Cagliari, 09042 Monserrato (Cagliari), Italy, <sup>c</sup>University of California Irvine, Irvine, CA 92697, USA, <sup>d</sup>Institute of Physics, Academy of Sciences of the Czech Republic, 182 21, Czech Republic, <sup>e</sup>CERN, CH-1211 Geneva, Switzerland, <sup>f</sup>Cornell University, Ithaca, NY 14853, USA, <sup>g</sup>University of Cyprus, Nicosia CY-1678, Cyprus, <sup>h</sup>Office of Science, U.S. Department of Energy, Washington, DC 20585, USA, <sup>i</sup>University College Dublin, Dublin 4, Ireland, <sup>j</sup>ETH, 8092 Zürich, Switzerland, <sup>k</sup>University of Fukui, Fukui City, Fukui Prefecture, Japan 910-0017, <sup>l</sup>Universidad Iberoamericana, Lomas de Santa Fe, México, C.P. 01219, Distrito Federal, <sup>m</sup>University of Iowa, Iowa City, IA 52242, USA, <sup>n</sup>Kinki University, Higashi-Osaka City, Japan 577-8502, <sup>o</sup>Kansas State University, Manhattan, KS 66506, USA, <sup>p</sup>Brookhaven National Laboratory, Upton, NY 11973, USA, <sup>q</sup>Queen Mary, University of London, London, E1 4NS, United Kingdom, <sup>r</sup>University of Melbourne, Victoria 3010, Australia, <sup>s</sup>Muons, Inc., Batavia, IL 60510, USA, <sup>t</sup>Nagasaki Institute of Applied Science, Nagasaki 851-0193, Japan, <sup>u</sup>National Research Nuclear University, Moscow 115409, Russia, <sup>v</sup>Northwestern University, Evanston, IL 60208, USA, <sup>w</sup>University of Notre Dame, Notre Dame, IN 46556, USA, <sup>x</sup>Universidad de Oviedo, E-33007 Oviedo, Spain, <sup>y</sup>CNRS-IN2P3, Paris, F-75205 France, <sup>z</sup>Universidad Tecnica Federico Santa Maria, 110v Valparaíso, Chile, <sup>aa</sup>The University of Jordan, Amman 11942, Jordan, <sup>bb</sup>Universite catholique de Louvain, 1348 Louvain-La-Neuve, Belgium, <sup>cc</sup>University of Zürich, 8006 Zürich, Switzerland, <sup>dd</sup>Massachusetts General Hospital, Boston, MA 02114 USA, <sup>ee</sup>Harvard Medical School, Boston, MA 02114 USA, <sup>ff</sup>Hampton University, Hampton, VA 23668, USA, <sup>gg</sup>Los Alamos National Laboratory, Los Alamos, NM 87544, USA, <sup>hh</sup>Università degli Studi di Napoli Federico I, I-80138 Napoli, Italy

## II. THE CDF II DETECTOR

The CDF II detector is described in detail elsewhere [12]. The detector is cylindrically symmetric around the Tevatron beam line. Tracking detectors are installed around the interaction point to reconstruct the trajectories of charged particles (tracks). The tracking systems are located within a superconducting solenoid that produces a 1.4 T magnetic field aligned with the  $p\bar{p}$  beams. Around the outside of the solenoid, calorimeter modules arranged in a projective-tower geometry measure the energies of charged and neutral particles. Drift chambers outside the calorimeter are used to detect muons, which typically deposit little energy in the calorimeter.

The central outer tracker (COT) is a 3.1 m long open-cell drift chamber that has 96 measurement layers in the region between 0.40 and 1.37 m from the beam axis, providing full track coverage in the pseudorapidity region  $|\eta| < 1.0$ . Sense wires are arranged in eight alternating axial and  $\pm 2^\circ$  stereo “superlayers” with 12 wires each. The position resolution of a single drift-time measurement is about 140  $\mu\text{m}$ . A five-layer double-sided silicon microstrip detector (SVX) covers the region between 2.5 to 11 cm from the beam axis. Three separate SVX barrel modules along the beam line cover a length of 96 cm, approximately 90% of the luminous beam interaction region. Three of the five layers combine an  $r$ - $\phi$  measurement on one side and a  $90^\circ$  stereo measurement on the other, and the remaining two layers combine an  $r$ - $\phi$  measurement with a small angle ( $\pm 1.2^\circ$ ) stereo measurement. The typical silicon hit resolution is 11  $\mu\text{m}$ . An intermediate silicon layers at a radius of 22 cm from the beam line in the central region links tracks in the COT to hits in the SVX. The fiducial range of the silicon detector extends to pseudorapidity magnitude  $|\eta| < 2.0$ .

Calorimeter modules are located outside the central tracking volume and solenoid. The inner electromagnetic layers consist of lead sheets interspersed with scintillators, while the outer hadronic layers consist of scintillators sandwiched between steel sheets. The calorimeter is split between central barrel ( $|\eta| < 1.1$ ) and forward end-plug ( $1.1 < |\eta| < 3.6$ ) sections. Individual towers in the central barrel subtend 0.1 in  $|\eta|$  and  $15^\circ$  in  $\phi$ . The sizes of the towers in the end plug calorimeter vary with  $|\eta|$ , subtending 0.1 in  $|\eta|$  and  $7.5^\circ$  in  $\phi$  at  $|\eta| = 1.1$ , and 0.5 in  $|\eta|$  and  $15^\circ$  in  $\phi$  at  $|\eta| = 3.6$ . The energy resolution in the electromagnetic calorimeters is  $14\%/\sqrt{E_T}$  in the central barrel and  $16\%/\sqrt{E} \oplus 1\%$  in the forward end-plug section, with the energies in units of GeV. The single-particle energy resolution in the hadronic calorimeters, measured using pions, ranges from  $75\%/\sqrt{E}$  in the central barrel to  $80\%/\sqrt{E} \oplus 5\%$  in the forward end-plug section, with the energies expressed in units of GeV.

The hadronization of quarks and gluons produced in the interaction leads to collimated groups of high-momentum particles called jets. These jets, along with photons and electrons, leave isolated energy deposits in contiguous groups of calorimeter towers, which can be summed together into an energy cluster. Electrons and photons are identified as isolated, mostly electromagnetic clusters, and quality requirements may be placed on the presence of a high- $p_T$  track ge-

ometrically matched to the cluster to more accurately identify electrons. Jets are identified as electromagnetic and hadronic clusters with the combined electromagnetic fraction  $E_{EM}/E_{total} = E_{EM}/(E_{EM} + E_{had}) < 0.9$ , clustered using the JETCLU cone algorithm [13] with a fixed cone size of  $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$ .

Outside the calorimeters, drift chambers detect muons. A four-layer stack of planar drift chambers detect muons with  $p_T > 1.4 \text{ GeV}/c$ , and another four layers of drift chambers behind 60 cm of steel detect muons with  $p_T > 2.0 \text{ GeV}/c$ . Both systems cover a region of  $|\eta| < 0.6$ , though they have different structure and their geometrical coverages do not overlap exactly. Muons in the region between  $0.6 < |\eta| < 1.0$  pass through at least four drift layers arranged within a conic section outside of the central calorimeter. Muons are identified as either COT tracks that extrapolate to hits in the muon detectors, or as isolated tracks unmatched to hits in the muon detectors that satisfy tighter tracking-quality requirements and extrapolate to calorimeter energy depositions consistent with a minimum ionizing particle.

## III. DATA SET AND EVENT SELECTION

We analyze the full data set of  $p\bar{p}$  collisions collected by the CDF II detector. We require events to be collected from periods when the tracking systems, calorimeters, muon detectors were all functioning properly, corresponding to an integrated luminosity of  $8.9 \text{ fb}^{-1}$ . Events are selected via a number of high- $E_T$  electron and high- $p_T$  muon online event-selection requirements (triggers). The majority of these triggers require at least one electron (muon) with  $E_T > 18 \text{ GeV}$  ( $p_T > 18 \text{ GeV}/c$ ). We require the events to contain two electrons (muons) with  $E_T > 20 \text{ GeV}$  ( $p_T > 20 \text{ GeV}/c$ ) and determine the trigger selection and event reconstruction efficiencies by comparing the number of data and simulated  $Z \rightarrow \ell\ell$  events containing exactly one jet with  $E_T > 20 \text{ GeV}$ .

For the final analysis, we select events with at least two leptons, and two or more jets. In the unlikely case that more than two charged leptons are reconstructed, we select the two leptons with highest  $p_T$ . In addition to the  $p_T$  requirements on the leptons, we require leptons associated with well-reconstructed tracks (central electrons,  $|\eta| < 1$ , and all muons) to be of opposite charge, and a reconstructed dilepton invariant mass,  $m_{\ell\ell}$ , consistent with the mass of the  $Z$  boson,  $76 < m_{\ell\ell} < 106 \text{ GeV}/c^2$ . We require both leading- $E_T$  jets to have  $E_T > 25 \text{ GeV}$  and  $|\eta| < 2.0$ , and to be spatially separated from the reconstructed leptons ( $\Delta R > 0.4$ ). Additionally, the two jets must be separated by  $\Delta R > 0.7$ . Finally, as the final state should contain no particles that are not reconstructed in the detector, we also require that the missing transverse energy,  $\cancel{E}_T$  [14], is less than 20 GeV.

After this selection, three major sources of background contribute. The dominant background comes from the production of a  $Z$  boson, decaying to an  $e^+e^-$  or  $\mu^+\mu^-$  pair, in association with two jets. Simulated events generated using ALPGEN [15], interfaced with PYTHIA [16] for showering, are used to estimate this background. The production cross

sections for  $Z + b\bar{b}$  processes are normalized to experimental measurements [17].

Another significant background results from jets misidentified as leptons. The contributions from these lepton *fakes* are estimated using data-driven methods. For muons, we use events with same-sign muon pairs (rather than opposite-sign) that otherwise satisfy all event selection requirements. For electrons, we derive a misidentification rate representing the likelihood for a jet to be misidentified as an electron, as a function of jet  $E_T$  and  $\eta$ , using jet-triggered data with minimal contributions from events with electrons. This rate is then applied to all possible electron-jet pairs in events from the high- $p_T$  electron data set, where the jet is then treated as a second electron, and the event selection requirements are otherwise applied normally.

While the requirement to have low  $\cancel{E}_T$  reduces its total contribution, top-quark-pair production, where each top quark decays into a leptonic final state ( $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow \ell^+\nu_\ell b\ell^-\bar{\nu}_\ell\bar{b}$ ), contributes events to the final event sample, especially in the heavy-flavor-tagged channel. We estimate  $t\bar{t}$  contributions using PYTHIA with  $\sigma_{t\bar{t}} = 7.5$  pb and  $m_t = 172.5$  GeV/ $c^2$ . Finally, ZW and ZZ signal events are also modeled using PYTHIA. The predicted and observed numbers of events are shown in Table I.

#### IV. JET ENERGY CALIBRATION

The energies of jets, measured in the calorimeter, are corrected to account for a number of effects that distort the true jet energy. These effects include changes in calorimeter performance as a function of  $|\eta|$  and time, contributions from multiple  $p\bar{p}$  interactions per beam crossing (pileup), contributions from the other partons in the interacting proton and antiproton (underlying event), the non-linear response of the calorimeter, and energy radiated outside of the jet cone. These jet energy scale (JES) corrections are described in detail in Ref. [18].

These energy corrections, however, do not attempt to account for potential differences in the modeled calorimeter response to jets originating from quarks and gluons. For example, the largest corrections modify the energy scale of the jets to more accurately match that of the initial parton energies and their resulting particle jets, and are derived using PYTHIA [16] dijet Monte Carlo (MC) simulations, but independently of the initiating type of parton. Differences in the response of gluon and quark jets between simulation and data lead to differences in the measured energies for these jets that are not covered by the previously-assigned systematic uncertainties on the JES [19].

We derive a data-driven correction for the response to quark and gluon jets in simulated events using two independent samples of jets with different compositions of quark- and gluon-initiated jets. In these samples, we derive a correction to the jet energy by balancing the transverse energy of the jets against particles of known transverse momentum. We use events where a jet recoils against a high- $E_T$  photon—a sample rich in quark jets (based on simulations modeled using PYTHIA)—and utilize the significant number of  $Z \rightarrow \ell^+\ell^- +$

jet events available in the full CDF data set, which have a larger fraction of gluon jets. The quark and gluon content of these two samples differ due to the difference in mass between the  $Z$  boson and the photon: because the  $Z$  boson mass is higher, the initial partons of the production process typically carry a higher fraction of the momentum of the proton than those involved in the production of high-energy photons. This leads to a difference in the quark and gluon content of these samples.

To derive a correction, we construct the *balance* of the jet with these more accurately measured reference particles:

$$K_{Z,\gamma} = (E_T^{\text{jet}}/p_T^{Z,\gamma}) - 1. \quad (1)$$

For unbiased measurements of the jet energy, the balance will equal zero. Samples of jets with non-zero balance could be corrected with a jet-energy correction factor of  $1/(K_{Z,\gamma} + 1)$ . However, rather than derive independent JES corrections for quark and gluon jets in data and simulation, we compare the balance in data and simulation and derive an additional correction to be applied to simulated jets, dependent upon whether these jets are matched to quarks or gluons. The correction to simulated quark-jet energies is  $+1.4 \pm 2.7\%$ , while the correction to gluon-jet energies is much larger:  $-7.9 \pm 4.4\%$ .

##### A. Data set and event selection

The data set and event selection for the  $Z$ -jet balancing sample largely follow those described in Sec. III. We require two leptons consistent with resulting from the decay of a  $Z$  boson and exactly one jet with  $E_T > 15$  GeV and no other jets with (uncorrected)  $E_T > 3$  GeV within  $|\eta| < 2.4$ . Additionally, we ensure that the  $Z$  boson and jet are azimuthally opposite (back-to-back) by requiring their azimuthal separation to exceed 2.8 radians, and we require that  $p_T^Z > 10$  GeV/ $c$ .

For the  $\gamma$ -jet balancing sample, we closely mirror the selection requirements described in Ref. [18]. We use events collected with an isolated-central-photon trigger over the same period of time as that of the high- $p_T$  lepton samples. We compare these data to PYTHIA simulations of both  $\gamma$ +jet production as well as dijet production, which also contributes to the  $\gamma$ -jet balancing sample.

To match the selection requirements of the isolated-central-photon trigger, we require  $E_T^\gamma > 27$  GeV and  $0.2 < |\eta_\gamma| < 0.6$  in both data and MC simulation. To decrease the contribution from dijet production, where a jet mimics the photon selection, we require the energy in the calorimeter and momentum in the tracking system contained within a cone of  $R = 0.4$  around the photon to be less than 1 GeV and 2 GeV/ $c$ , respectively. As in the  $Z$ -jet balancing sample, we require events to have exactly one measured jet with  $E_T > 15$  GeV and no other jets with (uncorrected)  $E_T > 3$  GeV within  $|\eta| < 2.4$ . We also demand the  $\Delta\phi$  between the jet and photon to be larger than 3.0 radians. We further reduce contamination by vetoing events with more than one reconstructed interaction

	All events	Heavy-flavor-tagged	Light-flavor-tagged	Untagged
$Z$ +jets	$8\,700 \pm 1\,100$	$93 \pm 14$	$1\,520 \pm 310$	$7\,100 \pm 970$
$Z$ + $b$ jets	$710 \pm 300$	$111 \pm 48$	$55 \pm 26$	$550 \pm 230$
$t\bar{t}$	$9.2 \pm 0.9$	$3.3 \pm 0.4$	$0.7 \pm 0.1$	$5.1 \pm 0.6$
Misidentified leptons	$330 \pm 170$	$4.8 \pm 2.4$	$41 \pm 20$	$280 \pm 140$
Predicted background	$9\,700 \pm 1\,200$	$212 \pm 55$	$1\,620 \pm 330$	$7\,900 \pm 1\,100$
$ZW + ZZ$	$313 \pm 29$	$12.8 \pm 1.6$	$89 \pm 12$	$212 \pm 22$
Total predicted events	$10\,000 \pm 1\,300$	$225 \pm 55$	$1\,710 \pm 330$	$8\,100 \pm 1\,100$
Data events	9 846	172	1 724	7 950

TABLE I: Predicted and observed numbers of events in the final event selection, where the numbers of events are rounded to the appropriate significant figures given the uncertainties. The uncertainties incorporate all systematic uncertainties summarized in Table VI and include an additional 10% uncertainty on the normalization of  $Z$ +jets events and a 6% uncertainty on the normalization of  $ZW + ZZ$  events, from the theoretical uncertainties on the production cross sections for those processes.

point, and by removing events with  $\cancel{E}_T/E_T^\gamma > 0.8$ , which likely contain activity from cosmic rays.

### B. Determination of corrections

We derive separate corrections for the quark- and gluon-jet energy scales in data and simulation using the  $Z$ -jet and  $\gamma$ -jet balancing samples in the following way. The balances of the  $Z$ -jet and  $\gamma$ -jet systems ( $K_Z$  and  $K_\gamma$ , respectively) can be expressed as linear combinations of independent quark- and gluon-balance variables ( $K_q$  and  $K_g$ , respectively), weighted by the sample-specific quark and gluon fractions ( $F_{Z,\gamma}^{q,g}$ )

$$K_Z = F_Z^q K_q + F_Z^g K_g = F_Z^q K_q + (1 - F_Z^q) K_g, \quad (2)$$

$$K_\gamma = F_\gamma^q K_q + F_\gamma^g K_g = F_\gamma^q K_q + (1 - F_\gamma^q) K_g. \quad (3)$$

Rewriting these expressions by solving for  $K_q$  and  $K_g$ , we find

$$K_q = \frac{1}{F_\gamma^q - F_Z^q} [(1 - F_Z^q) K_\gamma - (1 - F_\gamma^q) K_Z], \quad (4)$$

$$K_g = \frac{1}{F_\gamma^q - F_Z^q} [F_\gamma^q K_Z - F_Z^q K_\gamma]. \quad (5)$$

These expressions apply separately to experimental data and simulated data, yielding a different balance in data and Monte Carlo simulation ( $K^{\text{data}}$  and  $K^{\text{MC}}$ , respectively) and may include a dependence on the energy of the jet, with  $F_{Z,\gamma}^q = F_{Z,\gamma}^q(E_T^{\text{jet}})$  and consequently  $K = K(E_T^{\text{jet}})$ .

In order to solve for  $K_q$  and  $K_g$ , we require knowledge of  $K_{Z,\gamma}$  and  $F_{Z,\gamma}^q$ . We extract a value of  $K_{Z,\gamma}$  as a function of  $E_T^{\text{jet}}$  by constructing the balancing distribution, as defined in Eq. (1), in ranges of  $E_T^{\text{jet}}$ , and fit the core of the distribution with a Gaussian shape. We perform these fits separately in data and simulation and use the mean and uncertainty on the mean of the fitted Gaussian shape as the value of  $K_{Z,\gamma}(E_T^{\text{jet}})$  and its uncertainty. We use this estimation of the most probable value in order to avoid effects from a small fraction of

highly mismeasured jets, which may strongly bias the mean and median of the distribution.

The distributions of  $K_Z$  and  $K_\gamma$  in data and simulated data are shown in Fig. 1. Not only are jets measured poorly (the balance does not average to zero), but in the  $Z$ -jet balancing sample, largely dominated by gluon jets, there is significant disagreement between the correction factors for simulated jets and those in data. We do not see a similar disagreement in the  $\gamma$ -jet balancing sample, indicating that the simulation models the behavior of the jets in this quark-jet dominated sample accurately.

We determine  $F_{Z,\gamma}^q$  from simulation by matching jets to their originating partons. In the  $\gamma$ -jet balancing sample, the quark fraction is about 85% at  $E_T^{\text{jet}} \approx 30$  GeV and drops to about 71% at  $E_T^{\text{jet}} \approx 70$  GeV. In the  $Z$ -jet balancing sample, these fractions are roughly 38% and 49%, respectively, in the same  $E_T^{\text{jet}}$  regions. In data, it is not possible to directly match jets to their originating parton, and we must therefore rely on simulation to extract values of  $F_{Z,\gamma}^q(E_T^{\text{jet}})$ . Because we are trying to correct for discrepancies in the reconstruction of quark and gluon jets between data and simulation, we cannot simply use the simulation-derived  $F_{Z,\gamma}^q$  values from each jet  $E_T$  bin. Rather, we parametrize  $F_{Z,\gamma}^q$  from simulation as a function of  $p_T^{Z/\gamma}$  and determine  $F_{Z,\gamma}^{q,\text{data}}$  in each jet  $E_T$  bin of the data based on the  $p_T^Z$  or  $p_T^\gamma$  distribution in that bin.

Using Eqs. (4-5), we construct distributions of  $K_q$  and  $K_g$  as functions of the jet  $E_T$ , as shown in Fig. 2. We see good agreement between data and simulation in  $K_q$  but poorer agreement in  $K_g$ , where the data correction appears consistently lower than that for simulation. This suggests that the MC simulation is systematically overestimating gluon jet energies, relative to the data.

Using the distributions of  $K_q$  and  $K_g$ , we determine the corrections that need to be applied to simulated jets in order to best match the energy scale of the data. These MC simulation corrections are defined as  $(K_q^{\text{data}} + 1)/(K_q^{\text{MC}} + 1)$  for quark jets and  $(K_g^{\text{data}} + 1)/(K_g^{\text{MC}} + 1)$  for gluon jets, as shown in Fig. 3. Due to the photon trigger used to select the  $\gamma$ -jet balancing sample, reliable balancing information is not available for jets with transverse energies smaller than 27.5

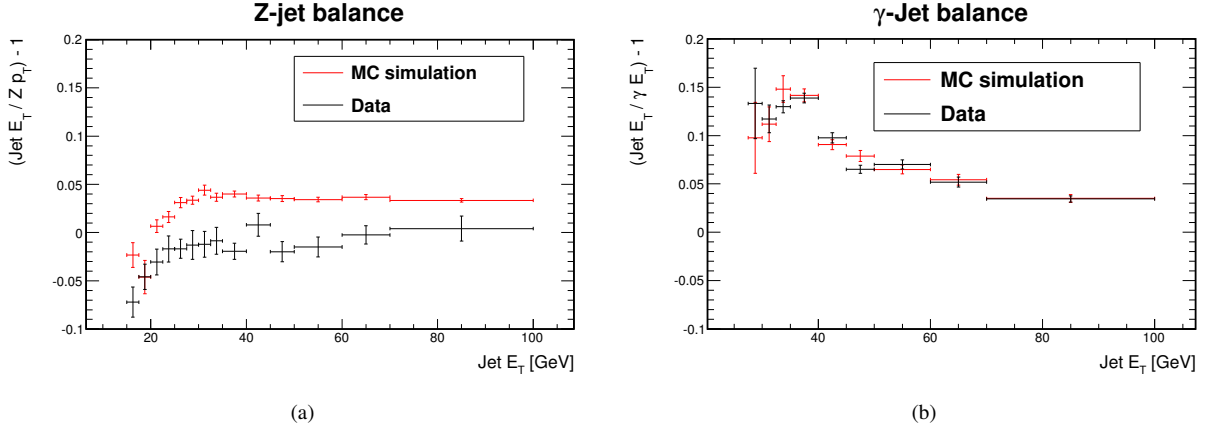


FIG. 1: Balancing distributions, (a)  $K_Z$  and (b)  $K_\gamma$ , in data and MC simulation as a function of  $E_T^{\text{jet}}$ . The uncertainties include solely the contribution from the fluctuations in the mean of a Gaussian fit to the balancing distributions in bins of  $E_T^{\text{jet}}$ .

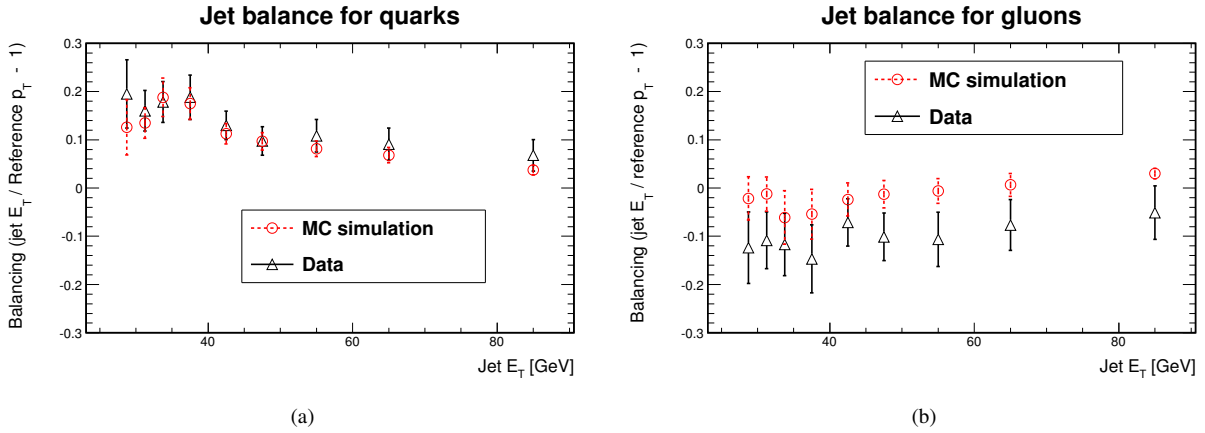


FIG. 2: Derived balancing variable for (a) quark jets,  $K_q$ , and (b) gluon jets,  $K_g$ , in data and MC simulation as a function of  $E_T^{\text{jet}}$ . The uncertainties on each point are from the uncertainties from the mean of the Gaussian fit and the uncertainties on the quark fractions, added in quadrature.

GeV in that sample, limiting the range of applicability of the corrections. Since we are interested in jets of energies extending down to 20 GeV, we extrapolate the quark-jet energy correction derived for jets with  $E_T > 27.5$  GeV to lower jet energies and use the  $Z$ -jet balancing sample to extract a gluon correction assuming this extrapolated quark correction.

The quark and gluon corrections' dependence on jet energy are accurately modeled by a constant for jets with  $E_T > 15$  GeV. Quark jet energies in simulation should be increased by approximately 1.4% to more accurately match the data, while gluon jet energies should be decreased by approximately 7.9%.

### C. Uncertainties on simulated jet energy corrections

We consider the following sources of uncertainty on the correction factors.

1. *Statistical uncertainty*: We use the standard deviation of the corrections to the simulation for each jet  $E_T$  bin around the assumed uniform correction function. This is an uncertainty of  $\pm 2.0\%$  for quark jet energies, and  $\pm 2.5\%$  for gluon jet energies (the uncertainties on quark- and gluon jet energy corrections are anticorrelated).
2.  $F_Z^q$ : We compare the distribution of a quark-gluon discriminant parameter (described in detail in Sec. V) in data and simulation, and fit the data distribution using quark and gluon templates from simulation. We take the average deviation of the value determined for this quark-gluon discriminant from the nominal MC simulation value as a systematic uncertainty on  $F^q$ , constant across jet  $E_T$ . This uncertainty is approximately 10%. Here we vary the calculated quark fraction in the data  $Z$ -jet balancing sample by  $\pm 10\%$  and recalculate the



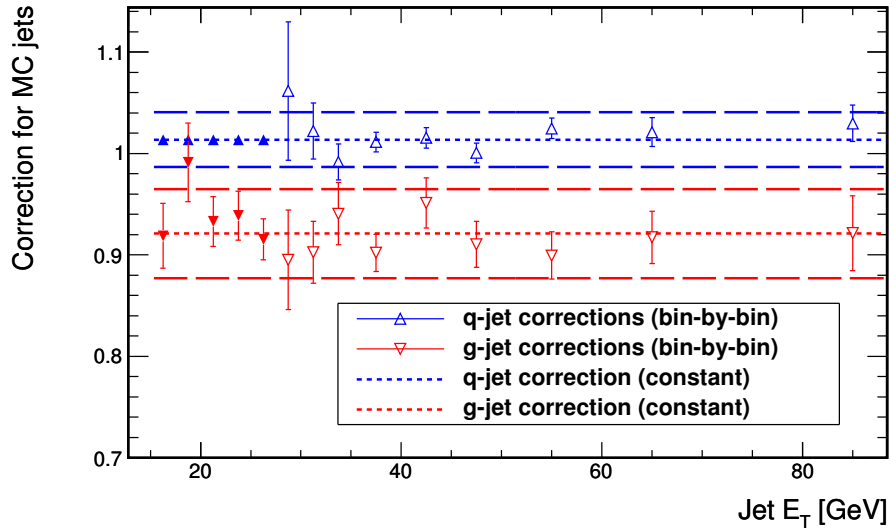


FIG. 3: Derived correction for simulated quark jets and gluon jets as a function of  $E_T^{\text{jet}}$ . The open triangles represent corrections derived using both  $\gamma$ -jet and  $Z$ -jet balancing samples, while the filled triangles represent the assumed uniform correction for quarks and the corresponding correction for gluons calculated from the  $Z$ -jet balancing sample alone. The uncertainties shown are statistical only. The short dashed lines are the fits of the correction to a constant across jet  $E_T$  (as opposed to the corrections in each bin of  $E_T^{\text{jet}}$ ), and the long dashed lines represent the total systematic uncertainty bands on that constant correction, further described in Sec. IV C.

corrections for quark and gluon jets. This translates to an uncertainty of  $\pm 0.6\%$  for quark jet energies and  $\mp 2.1\%$  for gluon jet energies.

3.  $F_\gamma^q$ : We follow a similar procedure of fitting the quark-gluon discriminant parameter in the  $\gamma$ -jet sample and obtain a similar uncertainty of  $\pm 10\%$  on the quark fraction. This translates to an uncertainty of  $\pm 1.8\%$  for quark-jet energies, and  $\mp 2.7\%$  for gluon jet energies.
4. *Low  $E_T$  extrapolation*: We check the dependence of the gluon jet energy corrections on the assumed quark jet corrections for low- $E_T$  jets by varying the quark jet  $E_T$  for these jets by  $\pm 2\%$ . We see a small change in the gluon energy corrections, which translates to  $\mp 0.4\%$  of the gluon jet energy.
5. *Number of interaction vertices dependence*: The  $\gamma$ -jet balancing sample incorporates a requirement on the number of reconstructed interaction vertices to reduce contamination from pileup. The  $Z$ -jet balancing sample does not have such a requirement, due to a much smaller background contribution and in order to retain as many events as possible. We check for any bias in the corrections resulting from the effect of this requirement by looking for any shift in the corrections when the requirement is placed on the  $Z$ -jet balancing sample. We see a change in the quark-jet energies of  $\pm 0.2\%$ , and the gluon jet energies of  $\mp 1.2\%$ .

The uncertainties are summarized in Table II. Because the corrections shift the energy response in the simulation to more

accurately match the data, the quark jet and gluon jet energy correction uncertainties are anticorrelated. The uncertainties are similar in magnitude to the default CDF jet energy scale uncertainties [18].

		Quark jets	Gluon jets
Jet energy correction		1.014	0.921
Uncertainty	Statistical	0.020	0.025
	$F_Q^{Z-\text{jet}}$	0.006	0.021
	$F_Q^{\gamma-\text{jet}}$	0.018	0.027
	Low $E_T$ extrapolation		0.004
	$N_{\text{vert}}$ difference	0.002	0.012
Total uncertainty		$\pm 0.027$	$\mp 0.044$

TABLE II: Summary of the additional jet energy corrections applied to simulated jets and their uncertainties.

## V. ARTIFICIAL NEURAL NETWORK QUARK-TO-GLUON DISCRIMINANT

In this analysis, we search for two high- $p_T$  leptons from the decay of a  $Z$  boson and two jets from a  $W \rightarrow qq'$  or  $Z \rightarrow q\bar{q}$  decay. Thus, the two signal jets are quark jets. Conversely, the dominant background, two jets produced in association with a  $Z \rightarrow \ell^+\ell^-$  decay, contains a significant fraction of gluon jets (of the order of 50%). The ability to separate quark jets from

gluon jets is therefore useful for increasing sensitivity to  $ZW$  and  $ZZ$  production.

For a given energy, gluon jets, due to their higher color charge, tend to feature a higher particle multiplicity and be spatially broader in the detector than light-quark ( $u$ ,  $d$ , and  $s$ ) jets. We attempt to quantify the spatial spread of jets using a collection of artificial neural-networks (NNs) trained to separate gluon jets from light-flavor quark jets. We refer to the output of the final NN as the jet quark-to-gluon discriminant value (or jet QG value). We calibrate the response of the final NN in MC simulation to match the response in data based on a  $W \rightarrow \ell\nu + 1$  jet event sample. The tagging efficiency and mistag rate associated to a requirement on the jet QG value are obtained from two independent event samples:  $W \rightarrow \ell\nu + 2$  jets events, which are representative of the  $Z +$  jets background; and  $t\bar{t} \rightarrow b\bar{b}\ell\nu q\bar{q}'$  events, which contain two non-heavy-flavor jets from the hadronic decay of a  $W$  boson, similar to the diboson signal.

### A. Jet QG definition

A total of three NNs contribute to the final QG discriminant. The initial two networks separate quark and gluon jets by exploiting distinctive features in the distribution of energies reconstructed in calorimeter towers and momenta of charged particles reconstructed in the tracking chambers. Thus, every jet is assigned a *tower* NN value and *track* NN value, which are the outputs of these networks. These two NN values are then used as inputs to a third NN.

Each of the NNs is trained using simulated samples of jets matched to either a light-flavor-quark or gluon with  $p_T > 20$  GeV/ $c$  within  $\Delta R = 0.4$  of the center of the jet and further requiring that no additional partons with transverse momenta exceeding 8 GeV/ $c$  are present within  $\Delta R = 0.7$ . These jets are selected from a  $Z \rightarrow \mu^+\mu^- + 2$  parton ALPGEN sample, interfaced with PYTHIA showering. Each NN is a feed-forward multilayer perceptron with a hyperbolic-tangent-like response function [20]. The networks are trained on 100 000 quark and gluon jets and tested for biases in overtraining on samples containing 500 000 quark and gluon jets. Gluon-jet distributions are reweighted to match the  $E_T$  and  $\eta$  distributions of the quark jets to remove any discrimination power coming solely from these variables.

For each jet we obtain a list of the calorimeter towers within a cone of  $\Delta R = 0.7$ . Each tower has a location coordinate,  $(\eta, \phi)$ , and energy deposition  $E$  associated with it. We construct a distribution of the distance,  $\Delta R$ , between all pairs of towers within the jet and weight each tower pair by its relevance in terms of energy to obtain a distribution that characterizes the spatial spread of the energy within each jet. The weight is given by

$$\frac{E_i E_j}{0.5 [(\Sigma E)^2 - \Sigma E^2]},$$

where  $E_i$  and  $E_j$  are the energies detected in the two towers of the pair,  $\Sigma E$  is the sum of the energy in all towers within a cone of  $\Delta R = 0.7$  around the jet, and  $\Sigma E^2$  is the sum of the

square of the energies of each tower in that same cone. The denominator is chosen to normalize the sum of the weights of all tower pairs to unity. We sample this distribution in 56 intervals (bins) of size  $\Delta R_{\text{bin}} = 0.025$  for  $0.0 < \Delta R < 1.4$ , where the contents of the first three bins are empty due to the segmentation of the calorimeter. Typical distributions of the weighted  $\Delta R$  between tower pairs for quark and gluon jets are shown in Fig. 4, using a larger bin size. The outputs of the tower NN for quark and gluon jets using the training and testing samples are shown in Fig. 5.

We follow a similar prescription using tracks within a cone of  $\Delta R = 0.7$  around each jet, using the tracks' locations in  $(\eta, \phi)$  (with respect to the primary vertex) and momenta  $p$  to obtain a distribution of the distance between pairs of tracks (in  $\Delta R$ ), with each pair weighted by the momentum carried by that pair, or

$$\frac{p_i p_j}{0.5 [(\Sigma p)^2 - \Sigma p^2]},$$

where  $p_i$  and  $p_j$  are the magnitude of the momenta of the charged particles in the pair,  $\Sigma p$  is the scalar sum of the momenta carried by all charged particles within a cone of  $\Delta R = 0.7$  around the jet, and  $\Sigma p^2$  is the sum of the square of the momenta of each charged particle within that same cone. We require all contributing charged particles to come from the primary vertex and have  $p_T > 0.4$  GeV/ $c$ . We split the  $\Delta R$  between track pairs distribution into the same 56 intervals (bins) as used in the tower NN, and the content of each bin is used as an input into the track NN.

Typical distributions of  $\Delta R$  between track pairs for quark and gluon jets are shown in Fig. 4. Light-flavor quark jets tend to peak at low  $\Delta R$ , indicating that they are rather collimated, while gluon jets tend to have a higher mean-valued  $\Delta R$  distribution. The bin contents of these  $\Delta R$  distributions are used as inputs into NNs that discriminate between quark and gluon jets.

The outputs of the track NN for quark and gluon jets using the training and testing samples are shown in Fig. 5. Higher NN scores indicate jets that are more quark-like. We see good performance in both the tower and track NNs. The cusps in the track NN distribution are associated to jets containing only two charged-particle tracks located inside a cone of  $\Delta R = 0.7$ , and thus have only one nonzero bin in their distributions of  $\Delta R$  between track pairs.

The final NN uses both the tower and track NN values as inputs, along with other jet variables that provide discrimination power between quark jets and gluons: the ratio of  $\Sigma E$  associated to towers within a cone of  $\Delta R = 0.4$  to  $\Sigma E$  associated to towers within a cone of  $\Delta R = 0.7$ ; the ratio of  $\Sigma p$  associated to charged particles within a cone of  $\Delta R = 0.4$  to  $\Sigma p$  associated to charged particles within a cone of  $\Delta R = 0.7$ ; the number of contributing towers with nonzero energy in cones of  $\Delta R = 0.4$  and  $0.7$ ; the number of contributing charged-particle tracks in cones of  $\Delta R = 0.4$  and  $0.7$ ; and the jet EM fraction. Additionally, other variables that affect the shape of the  $\Delta R$  distributions, independent of whether the jet originates from a quark or gluon, are included: the jet  $E_T$ ; the jet  $\eta$ ; and, the number of reconstructed interaction vertices in

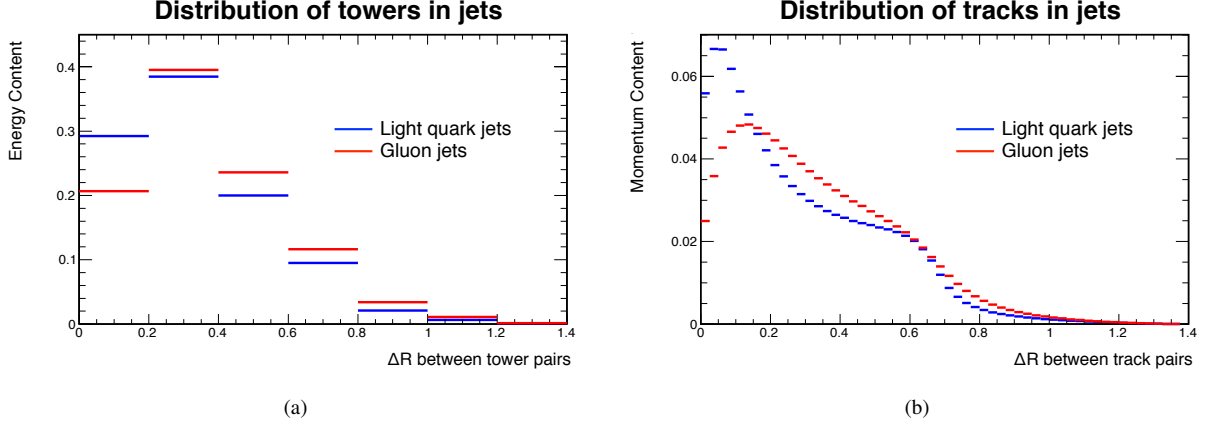


FIG. 4: Typical distributions of energy (momentum) content of jets as a function of the  $\Delta R$  (a) between pairs of towers and (b) between pairs of tracks in light-flavor quark and gluon jets in simulation.

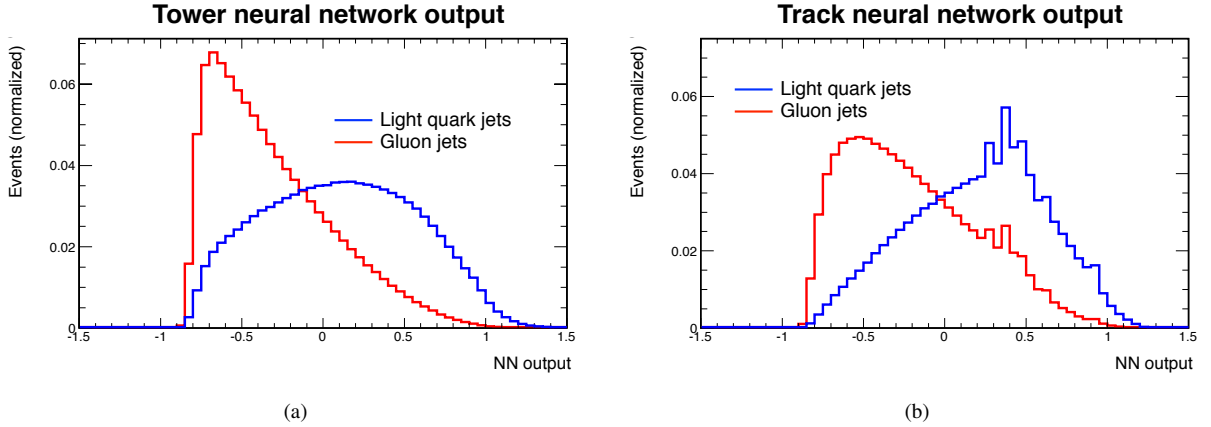


FIG. 5: Distributions of the outputs of the NNs processing (a) tower information and (b) track information in light-flavor quark and gluon jets in simulation. Higher NN scores indicate jets that are more quark-like.

the event. The output of this final NN is shown in Fig. 6 for light-flavor-quark and gluon jets using the training and testing samples. In simulated jets, we see significant separation between quark and gluon jets using this discriminant.

### B. Jet QG calibration and performance

The response of the NN quark-to-gluon discriminant may differ between data and MC simulation, especially since uncorrected tower energies are used in the construction of the tower NN. Since the signal and most backgrounds are modeled with simulated data, we calibrate the simulation response to match the response in data. We use a control region of independent events with a jet composition similar to that of the final state,  $W \rightarrow \ell\nu + 1$  jet events. We then validate the calibration and establish uncertainties on the modeling using control samples of data with features similar to the signal and samples enriched in the dominant backgrounds:  $t\bar{t}$  decays in

lepton+jets final states and  $W \rightarrow \ell\nu + 2$  jet events, respectively.

To form the  $W + 1$  jet calibration sample, we choose data collected with the standard high- $E_T$  ( $p_T$ ) central electron (muon) triggers and select events with exactly one central ( $|\eta| < 1.0$ ) electron (muon) with  $E_T$  ( $p_T$ )  $> 20$  GeV/ $c$ . To select events consistent with a  $W \rightarrow \ell\nu$  decay, we also require a significant missing transverse energy,  $\cancel{E}_T > 25$  GeV, and a reconstructed transverse mass [21] consistent with leptonic  $W$  boson decays,  $m_T > 25$  GeV/ $c^2$ . To further suppress any contributions from multijet events where a jet mimics the lepton +  $\cancel{E}_T$  signature, we require that the  $\cancel{E}_T$  is not aligned with any reconstructed jet ( $\Delta\phi(\cancel{E}_T, \text{jet}) > 0.2$  radians) and that the  $\cancel{E}_T$ -significance—a dimensionless quantity comparing the observed  $\cancel{E}_T$  against the energy resolution of jets, soft unclustered particles, and the event topology (see Ref. [5])—be larger than four (one) for events with electrons (muons). We also require that the events in this calibration sample have exactly one jet with  $E_T > 20$  GeV and  $|\eta| < 2.0$ .

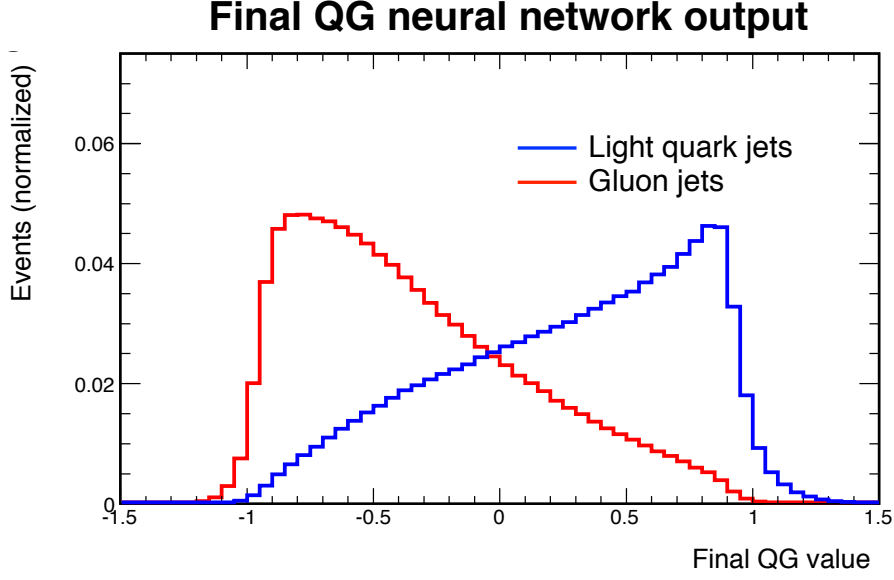


FIG. 6: Distribution of the output of the final NN for light-flavor-quark and gluon jets in simulation. Higher NN scores indicate jets that are more quark-like.

We consider various processes that contribute to this sample, listed in Table III, and model them using a combination of the PYTHIA, ALPGEN, and MADGRAPH [22] event generators interfaced with PYTHIA for showering. The dominant contribution is  $W \rightarrow \ell\nu$  production in association with one jet, which is modeled using ALPGEN. As we are largely concerned with the agreement in shapes between data and simulation, we scale the simulation distributions to match the data. Additionally, we reweight the simulation to match the jet  $E_T$  and  $\eta$  distributions in data to remove these variables as a possible causes for mismodeling of the jet QG value.

We observe poor modeling of the tower NN values, where the jets in data appear more gluon-like than those in simulated events. The fact that jets in data appear more spatially spread than jets in simulation is consistent with the observed differences in jet energy scales for data and simulation, described in Sec. IV: the fraction of the jet energy contained within a cone of  $\Delta R = 0.4$  is higher in simulated gluon jets than in gluon jets in data. We correct for these discrepancies by applying a linear shift to the tower NN values observed in simulation in order to match with data using the  $W + 1$  jet sample. We apply different linear shifts for jets in the central and plug calorimeters, and for jets in events with different levels of pileup. We apply further corrections to the response of the final NN to more accurately match the correlations of these calibrated tower-NN values with other jet quantities: the number of towers in the jets and the ratio of  $\Sigma E$  in a cone of  $\Delta R = 0.4$  to  $\Sigma E$  in a cone of  $\Delta R = 0.7$ . The modeling is more accurate in the track NN than in the tower NN, though we still introduce a similar linear shift in simulated track NN values to more accurately match data. The calibrated variables are input directly into the final NN, without retraining

the network.

We further validate the response of the jet QG value by comparing data and MC simulation in a  $W \rightarrow \ell\nu + 2$  jets event sample and in an event sample dominated by  $t\bar{t}$  production where two quark jets originate from the hadronic decay of a  $W$  boson. Table IV summarizes the requirements used to select these two samples: the  $W + 2$  jet sample is similar to the previously described  $W + 1$  jet sample, except for modified jet selections to match those used in the signal region of the  $ZV \rightarrow \ell\ell jj$  search. The  $t\bar{t}$  selection eschews the  $\cancel{E}_T$ -significance and  $m_T$  requirements, used to reduce multijet backgrounds, in favor of requirements on the minimum scalar sum of transverse quantities (jets'  $E_T$ ,  $\cancel{E}_T$ , and the charged lepton's  $p_T$ ) in the event, which is effective in removing both multijet and  $W +$  jets backgrounds. Because we are interested in selecting the two jets in the  $t\bar{t}$  candidate events that come from the decay of a  $W$  boson, as opposed to the  $b$  jets produced in the  $t \rightarrow Wb$  decays, we make use of the jet- $b$ ness tagger [23]. This multivariate  $b$  jet identification algorithm exploits properties of individual charged-particle tracks within a jet, looking at properties characteristic of charged particles originating from  $B$ -hadron decays. The final score, the output of a NN discriminant that ranges between  $-1$  and  $1$ , is called the *jet bness*, where higher scores identify jets that are more likely to originate from  $B$ -hadron decays. We classify the two jets with the highest *bness* scores in the event as the two  $b$  jets, and the remaining two jets as those resulting from a  $W \rightarrow q\bar{q}'$  decay.

Because we are looking for jet QG shape differences between data and simulation that induce acceptance uncertainties when a requirement on the jet QG value is applied, we scale the number of  $W +$  jet events in simulation to match the

	$W$ +jets selection	$t\bar{t}$ selection
$W$ +jets	$21\,500 \pm 2\,200$	$38.7 \pm 3.9$
$W$ + $b$ jets	$940 \pm 380$	$13.8 \pm 5.5$
$Z$ +jets	$1\,250 \pm 130$	$3.1 \pm 0.3$
$Z$ + $b$ jets	$86 \pm 34$	$1.4 \pm 0.6$
$WW + WZ$	$1\,386 \pm 83$	$5.9 \pm 0.4$
Single top-quark	$767 \pm 77$	$19.6 \pm 2.0$
$t\bar{t}$	$1\,378 \pm 83$	$469 \pm 28$
$t\bar{t}$ ( $b$ jets)		$108 \pm 7$
$t\bar{t}$ ( $q$ jets)		$361 \pm 22$
Total expected	$27\,300 \pm 2\,200$	$551 \pm 30$
Data	27 319	579

TABLE III: Number of events in the  $W$ +2 jets and  $t\bar{t}$  lepton+jets region, showing only the uncertainties assigned on the normalization of each sample. The  $W$ + jets samples are rescaled to match data in the number of events observed after the  $W$ + jets selection. The distinction between  $b$  and  $q$  jets in the  $t\bar{t}$  sample refers to the lower two bness jets: events where both jets are matched to non- $b$ -quark jets are labeled  $q$  jets, while if one of the jets is matched to a  $b$  jet, the event is labeled  $b$  jets.

$W$ +jets selection	$t\bar{t}$ selection
Central $e$ or $\mu$ , $p_T > 20$ GeV/ $c$ $\cancel{E}_T > 25$ GeV	
$\Delta\phi(\cancel{E}_T, \text{nearest jet}) > 0.4$ rad $\cancel{E}_T\text{-sig} > 4$ ( $e$ only) $m_T(W) > 25$ GeV/ $c^2$ ( $e$ only)	$\Delta\phi(\cancel{E}_T, \text{nearest jet}) > 0.2$ rad
$N_{jets}(E_T > 20 \text{ GeV}) = 2$	Sum $E_T > 300$ GeV $N_{jets}(E_T > 20 \text{ GeV}) = 4$ 2nd highest bness jet $b_{\text{ness}} > -0.5$ 2nd highest bness jets $E_T > 20$ GeV 2 lowest bness jets $E_T > 25$ GeV
1st/2nd jet $E_T > 25$ GeV	
Jet $ \eta  < 2.0$ $\Delta R$ between jets $> 0.7$	

TABLE IV: Summary of event selection requirements for the  $t\bar{t}$  lepton+jets selection and the  $W$ +2 jets selection, used to understand the modeling of events in the QG discriminant. Requirements in the center are shared requirements in the two samples.

yield observed in the  $W + 2$  jets data. The number of events in each sample is shown in Table III. The distributions of the maximum and minimum QG values of the two jets are shown in Figs. 7-8. We see fairly good modeling in the  $t\bar{t}$  sample, but poorer modeling in the  $W + 2$  jet sample where, even after calibrations, the jets in simulation appear more gluon-like than the jets in data. We account for this remaining discrepancy between data and simulation below.

We enhance the sensitivity to the signal when forming a *light-flavor-tagged* channel where the minimum jet QG value is greater than 0.0. We determine a probability for a quark jet to meet this requirement (efficiency), and for a gluon jet to be misidentified as a quark jet (mistag rate), with the  $t\bar{t}$  and  $W + 2$  jet samples. The efficiency measured in data,  $e_D$  is a function of the QG requirement,  $q$ , and may be expressed as

$$e_D(q) = \frac{e_{\text{raw}}(q) - s_m(q)m_{\text{MC}}(q)f_g}{1 - f_g}, \quad (6)$$

where  $e_{\text{raw}}$  is the fraction of data events passing the QG requirement;  $m_{\text{MC}}$  is the mistag rate for gluons, as measured in simulation;  $s_m$  is a scale factor on the mistag rate in simulated

jets to match the mistag rate measured in data; and  $f_g$  is the fraction of gluon jets in the sample. We can write a similar expression for the mistag rate from

$$m_D(q) = \frac{m_{\text{raw}}(q) - s_e(q)e_{\text{MC}}(q)f_q}{1 - f_q}, \quad (7)$$

where  $m_{\text{raw}}$  is the fraction of data events meeting the QG requirements;  $e_{\text{MC}}$  is the efficiency for quarks to pass the requirement, as measured in simulation;  $s_e$  is a scale factor on the efficiency in simulated jets to match the mistag rate measured in data; and  $f_q$  is the fraction of quark jets in the sample. Squared uncertainties on these quantities may be expressed as

$$\begin{aligned} \sigma_e^2(q) = & \frac{1}{(1 - f_g)^2} \left[ \frac{e_{\text{raw}}(1 - e_{\text{raw}})}{N_D} + (\sigma_m f_g)^2 \right] \\ & + \sum_X \frac{\sigma_X^2}{[N_{\text{MC}}(1 - f_g)]^2} \times \\ & [(e + s_m m_{\text{MC}})(f_g - f_g^X) + f_q^X (e_{\text{MC}} - e_X)]^2, \quad (8) \end{aligned}$$

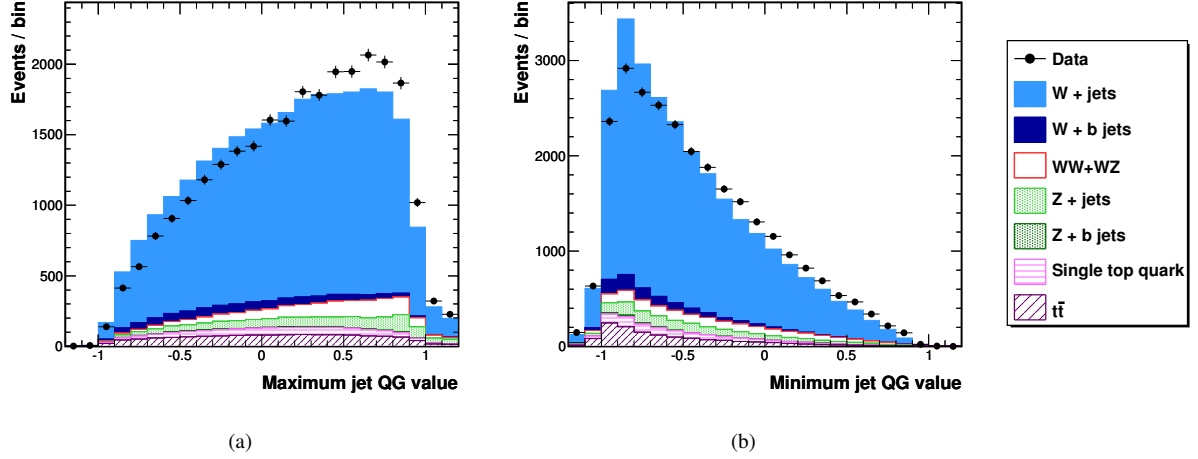


FIG. 7: Distribution of (a) the maximum and (b) minimum jet QG values of the two jets in the  $W + 2$  jet sample.

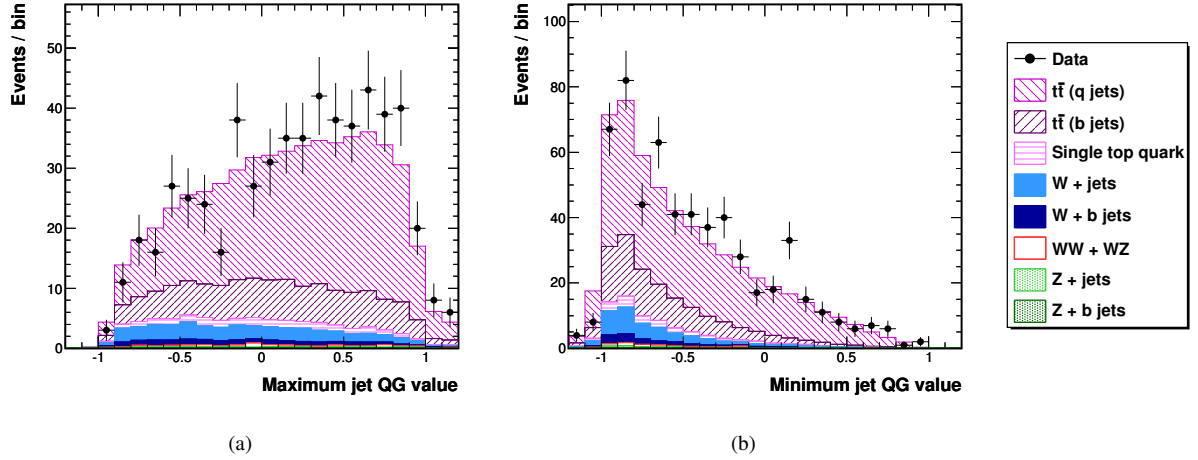


FIG. 8: Distribution of (a) the maximum and (b) minimum jet QG values of the two jets in the  $t\bar{t}$  sample. The distinction between  $q$  and  $b$  jets refers to the lower two  $b$ ness jets: events where both jets are matched to non- $b$  quark jets are labeled  $q$  jets, while if one of the jets is matched to a  $b$  jet, the event is labeled  $b$  jets.

where  $N_D$  and  $N_{MC}$  are the number of data and simulated events, respectively, and where the  $X$  represents the various simulated subsamples, and  $\sigma_m$  is the uncertainty on the mistag rate, which may be expressed in an analogous fashion. The uncertainty includes a statistical uncertainty on the data, uncertainties on the mistag rate and efficiency, and uncertainties on the relative difference in the contributions from the simulation. We take the uncertainties on the normalizations of the  $t\bar{t}$ , single top-quark, diboson,  $V$ +jets, and  $V + b$  jets to be 6%, 10%, 6%, 10%, and 40%, respectively, based on the uncertainties in their production cross sections.

We measure the efficiency in the  $t\bar{t}$  sample, where the fraction of gluon jets is small, and measure the mistag rate in the  $W + 2$  jets sample, where the gluon fraction is much larger

and similar to the fraction in the  $Z + 2$  jets signal region. The efficiency, mistag rate, and their uncertainties are determined using an iterative procedure. We first calculate the mistag rate in data assuming that the efficiency in data equals the efficiency in simulation. We then calculate the efficiency in data assuming that value for the mistag rate and proceed to update the mistag rate assuming the new value for the efficiency from data. We observe rapid convergence on robust values for the efficiency and mistag rate. Table V shows the efficiency and mistag rate for the given requirement of minimum QG  $> 0.0$ , measured in both data and MC simulation. The simulation underestimates the rate for quark jets to meet the jet QG requirement, while correctly predicting the observed mistag rate.

We implement a correction to the MC simulation by vary-

	MC	Data	MC revised jet QG requirement ( $-1\sigma, \text{nom.}, +1\sigma$ )
Efficiency	0.241	$0.295 \pm 0.034$	( $-0.0325, -0.09, -0.14$ )
Mistag rate	0.088	$0.087 \pm 0.027$	( $0.09, -0.0175, -0.11$ )

TABLE V: Efficiency and mistag rates for the chosen jet QG requirements, as evaluated in data and MC simulation, along with the necessary change in the jet QG threshold for the simulation to model the proper rates and the uncertainties on them.

ing the requirement on the minimum QG value in order to reproduce the efficiency and mistag rate observed in data. The uncertainties on these quantities are also obtained by varying the jet QG requirement. The alternate thresholds used for simulated quark and gluon jets are listed in Table V.

## VI. SIGNAL EXTRACTION AND RESULTS

We extract the number of signal events using a binned  $\chi^2$ -minimization fit to data, using the techniques described in Ref. [24]. We create histogram templates for both signal and background samples. The templates, along with the uncertainties we assign to their normalization in the fit procedure, are listed below.

1.  $ZV$  signal: We allow the normalization of the signal template to float unconstrained in the fit. We assume that each signal process contributes proportionally to its predicted SM cross section: 3.6 pb for  $ZW$  and 1.5 pb for  $ZZ$  [1].
2.  $Z$ +jets: This is the largest background. We allow its normalization to float in the fit, unconstrained.
3.  $Z+b$  jets: We constrain the normalization of this significant background to be within  $\pm 40\%$  of its nominal value.
4.  $t\bar{t}$ : We use a production cross section of  $\sigma_{t\bar{t}} = 7.5$  pb, and assign an uncertainty of 6.5% to the normalization of this template, based on the theoretical cross-section uncertainty [25].
5. Misidentified leptons: We use the method described in Sec. III to construct templates for the contribution from jets mimicking one or two leptons. We assign an uncertainty of 50% to the misidentification rate, based on studies using different trigger thresholds for the jet data used to obtain these rates.

We perform a simultaneous fit to data using independent templates for each of three channels. For events passing the basic signal selection requirements described in Sec. III, we first construct a heavy-flavor tag (HF-tag) channel composed of events passing a minimum jet  $b$ -ness requirement (jet  $b$ -ness  $> 0$ ), using the jet- $b$ -ness tagger [23]. For events failing this requirement, we then select events passing the minimum jet-QG value requirement described in Sec. V to form a light-flavor tag (LF-tag) channel. Events failing these requirement

are then placed in the third, untagged, channel, which has a lower signal fraction than the two tagged channels, but still includes a significant amount of signal due to the tight tagging requirements.

Additional systematic uncertainties on both the normalization and shapes of the templates used in the fit are also considered. We estimate uncertainties due to mismodeling between data and MC simulation in the jet energy scale (as described in Sec. IV C) and the jet energy resolution, the modeling of the tagging variables, and the lepton energy scale and resolution. Additional shape uncertainties on the  $Z$ +jets backgrounds are considered by increasing and decreasing the renormalization and factorization scale,  $Q^2$ , from the default value in the simulation of  $m_Z^2 + p_{T,Z}^2$ . We also consider the effect on the shape of the dijet invariant mass when increasing or decreasing initial- and final-state QCD radiation (ISR/FSR) in the  $ZV$  signal model. These systematic uncertainties, along with the normalization constraints described above, are treated as nuisance parameters in the fit, and are included in the  $\chi^2$ -minimization procedure [24]. They are summarized in Table VI.

Figure 9 shows the dijet mass distributions in data with the fit results overlaid. Table VII shows the number of events of each class determined by the fit. We fit for approximately 50% of the expected signal normalization and observe good agreement between data and simulation in the final fit for each of the three channels, with a total  $\chi^2/\text{d.o.f.} = 59.8/55$ .

We do not see significant evidence for  $ZW + ZZ$  production in this decay channel. Hence, we set upper limits on the production cross section using likelihood-ratio ordering [26], where we analyze the distribution of observed cross sections in pseudoexperiments generated with a variety of scale factors on the input signal cross section. When generating pseudoexperiments, we consider additional systematic uncertainties that affect the acceptance, assigning a 2% uncertainty on the signal template from limited knowledge of the from parton distribution functions, and 2.5% and 6% uncertainties due to the uncertainties on the lepton-scale-factor and integrated luminosity, respectively. The set of input cross sections in the pseudoexperiments range from 0.0 to 2.9 times the expected cross section, with a step size of 0.1.

Figure 10 shows the resulting confidence band. Using the  $1\sigma$  bands, we determine  $\sigma(p\bar{p} \rightarrow ZW + ZZ) = 2.5^{+2.0}_{-1.0}$  pb, compared to the standard model prediction of  $\sigma_{\text{SM}} = 5.1$  pb. We do not exclude the no-signal hypothesis, and establish a limit of  $\sigma_{ZW+ZZ} < 6.1$  pb ( $1.25 \times \sigma_{\text{SM}}$ ) at the 95% C.L.

Source	Channel	ZV	Z+jets	Z+b jets	$t\bar{t}$	Mis-ID leptons
Cross section/norm.	All	Unconstr.	Unconstr.	$\pm 40\%$	$\pm 6.5\%$	$\pm 50\%$
Jet energy res.	HF-tag	$\pm 0.8\%$	$\pm 0.3\%$	$\pm 1.0\%$	$\pm 0.2\%$	
	LF-tag	$\pm 1.0\%$	$\pm 0.7\%$	$\pm 1.5\%$	$\pm 6.2\%$	
	Untagged	$\pm 0.6\%$	$\pm 0.9\%$	$\pm 0.7\%$	$\pm 1.1\%$	
Jet energy scale	HF-tag	$\pm 4.0\%$	$\pm 4.4\%$	$\pm 3.8\%$	$\pm 4.0\%$	
	LF-tag	$\pm 1.5\%$	$\pm 0.3\%$	$\pm 0.6\%$	$\pm 3.0\%$	
	Untagged	$\pm 1.9\%$	$\pm 5.7\%$	$\pm 3.8\%$	$\pm 1.9\%$	
$Q^2$	All	none	Shape only	Shape only	none	
ISR/FSR	All	Shape only	none	none	none	
bness tag	HF-tag	$\pm 7.8\%$	$\pm 7.8\%$	$\pm 9.2\%$	$\pm 7.6\%$	
	LF-tag	$\pm 0.2\%$	$\pm 0.0\%$	$\pm 1.2\%$	$\pm 2.8\%$	
	Untagged	$\pm 0.4\%$	$\pm 0.1\%$	$\pm 1.8\%$	$\pm 4.5\%$	
QG tag	LF-tag	$\pm 10\%$	$\pm 16\%$	$\pm 2.0\%$	$\pm 15\%$	
	Untagged	$\pm 4.3\%$	$\pm 3.5\%$	$\pm 2.0\%$	$\pm 2.0\%$	
Lepton energy scale	All	$\pm 0.5\%$	$\pm 0.5\%$	$\pm 0.5\%$	$\pm 1.5\%$	
Lepton energy res.	All	$\pm 0.1\%$	$\pm 0.1\%$	$\pm 0.0\%$	$\pm 2.7\%$	

TABLE VI: Summary of the systematic uncertainties considered in the fit of the dijet mass distribution. Uncertainties that change both the shape and rate of templates used in the fit are treated in a correlated fashion.

Process	$N_{events, \text{HF-tag}}$	$N_{events, \text{LF-tag}}$	$N_{events, \text{Untagged}}$
Z+jets	$91.9 \pm 8.3$	$1\,605 \pm 50$	$7\,200 \pm 600$
Z+b jets	$71 \pm 14$	$37 \pm 10$	$360 \pm 100$
$t\bar{t}$	$3.2 \pm 0.4$	$0.7 \pm 0.1$	$5.3 \pm 0.4$
Misidentified leptons	$4.6 \pm 2.3$	$39 \pm 20$	$270 \pm 140$
Total background	$171 \pm 14$	$1\,681 \pm 36$	$7\,840 \pm 600$
$ZW + ZZ$	$6.3 \pm 4.4$	$45 \pm 30$	$106 \pm 72$
Total events	$177 \pm 14$	$1\,726 \pm 40$	$7\,940 \pm 610$
Data events	172	1\,724	7\,950

TABLE VII: Number of events in each class from the best fit to the data.

## VII. SUMMARY

In this paper we describe a search for  $ZW$  and  $ZZ$  boson pair-production from a final state with two charged, high-transverse-momentum electrons or muons and two hadronic jets. We increase the sensitivity by tagging events with jets likely originating from heavy- and light-flavor quarks and classifying them in separate analysis channels using neural-network-based taggers. These taggers benefit from the large sample of events containing top quarks collected by CDF, allowing a data-driven estimate of the efficiency and mistag rates for jets passing tagging requirements. We also improve the modeling of the Monte Carlo simulations, especially those that describe the  $Z$ + jets background, by deriving and incorporating improved energy corrections for simulated jets to more accurately reproduce the phenomenology of jets originating from quarks and gluons in the data.

Using the full CDF Run II proton-antiproton collisions data set, which corresponds to an integrated luminosity of  $8.9 \text{ fb}^{-1}$ , we fit for the normalization of  $ZW, ZZ \rightarrow \ell^+ \ell^- + q\bar{q}'$  events using the dijet invariant mass distribution. We incorporate many of the systematic uncertainties associated with the modeling of signal and background processes as nuisance

parameters in the dijet mass fit. We measure a cross section of  $\sigma_{ZW+ZZ} = 2.5^{+2.0}_{-1.0} \text{ pb}$ , which is nonzero at the  $1.75\sigma$  level of significance. We also obtain a limit on the cross section of  $\sigma_{ZW+ZZ} < 6.1 \text{ pb}$  at the 95% C. L.

## Acknowledgments

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean World Class University Program, the National Research Foundation of Korea; the Science and Technology Facilities Council and the Royal Society, United Kingdom; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, and Programa Consolider-Ingenio



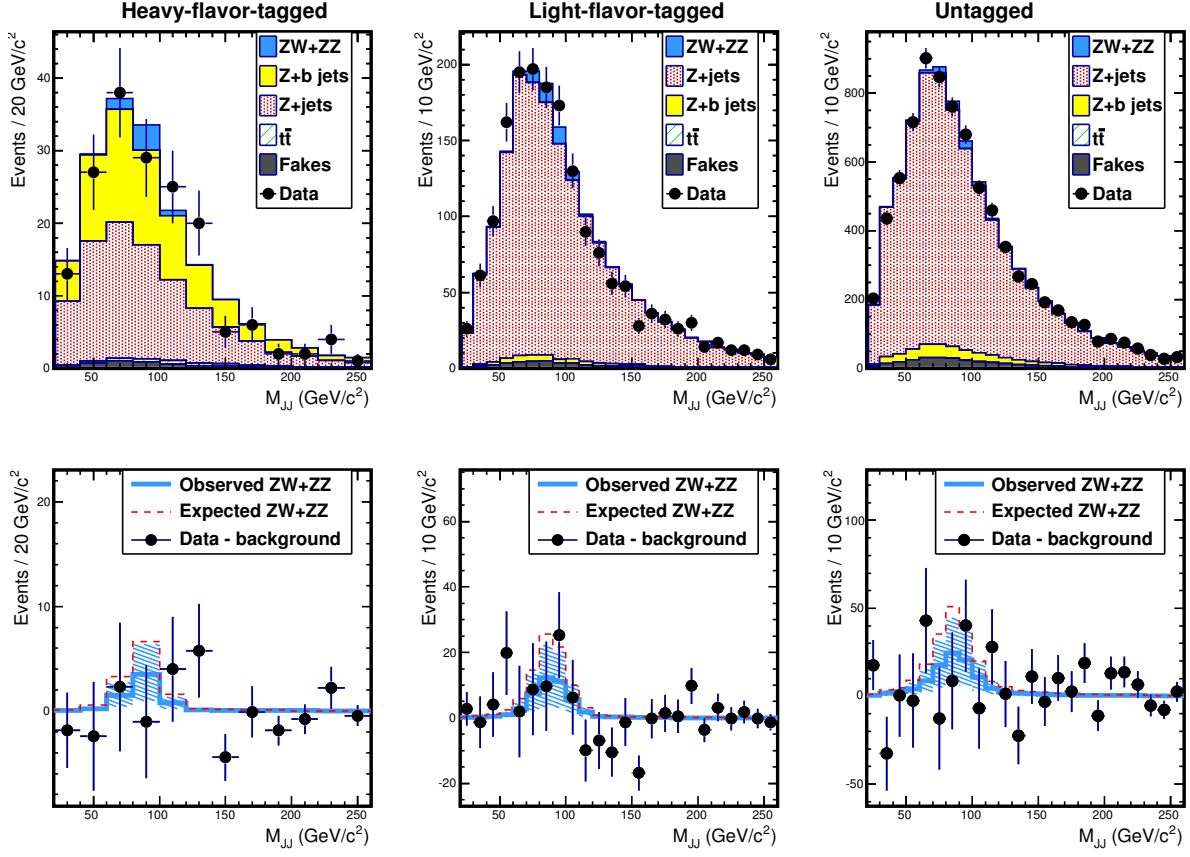


FIG. 9: Invariant dijet mass distributions with fit results overlaid for the  $ZW + ZZ$  process in the dilepton+dijet selection in the heavy-flavor-tagged channel (left panels), light-flavor-tagged channel (center panels), and untagged channel (right panels). The top row shows the output from the fit compared to the data, while the bottom row shows the background subtracted from data, compared to the expected (dashed line) and fitted (solid line, with uncertainties in bands) signal contributions.

2010, Spain; the Slovak R&D Agency; the Academy of Finland; the Australian Research Council (ARC); and the EU

community Marie Curie Fellowship Contract No. 302103.

[1] J. M. Campbell and R. K. Ellis, Phys. Rev. D **60**, 113006 (1999).  
[2] K. Hagiwara, R. Peccei, D. Zeppenfeld, and K. Hikasa, Nucl. Phys. **B282**, 253 (1987).  
[3] M. Kober, B. Koch, and M. Bleicher, Phys. Rev. D **76**, 125001 (2007).  
[4] E. J. Eichten, K. Lane, and A. Martin, Phys. Rev. Lett. **106**, 251803 (2011).  
[5] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. Lett. **103**, 091803 (2009).  
[6] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. Lett. **104**, 101801 (2010).  
[7] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. Lett. **106**, 171801 (2011).  
[8] V. M. Abazov et al. (D0 Collaboration), Phys. Rev. Lett. **108**, 181803 (2012).  
[9] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. D **85**, 012002 (2012).  
[10] W. Ketchum, Ph.D. thesis, University of Chicago. FERMILAB-

THESIS-2012-36 (2012).

[11] The coordinate system used to describe the detector has its origin located at the detector's center. The proton beam direction is defined as the  $+z$  direction. The polar angle,  $\theta$ , is measured from the origin with respect to the  $z$  axis, and  $\phi$  is the azimuthal angle. Pseudorapidity, transverse energy, and transverse momentum are defined as  $\eta = -\ln \tan(\theta/2)$ ,  $E_T = E \sin \theta$ , and  $p_T = p \sin \theta$ , respectively. The rectangular coordinates  $x$  and  $y$  point radially outward and vertically upward from the Tevatron ring, respectively.  
[12] A. Abulencia et al. (CDF Collaboration), J. Phys. G **34**, 2457 (2007).  
[13] F. Abe et al. (CDF Collaboration), Phys. Rev. D **45**, 1448 (1992).  
[14] We define the missing transverse momentum  $\vec{\cancel{E}}_T \equiv -\sum_i E_T^i \mathbf{n}_i$ , where  $\mathbf{n}_i$  is the unit vector in the azimuthal plane that points from the beamline to the  $i$ th calorimeter tower. We call the magnitude of this vector the missing transverse energy,  $\cancel{E}_T$ .

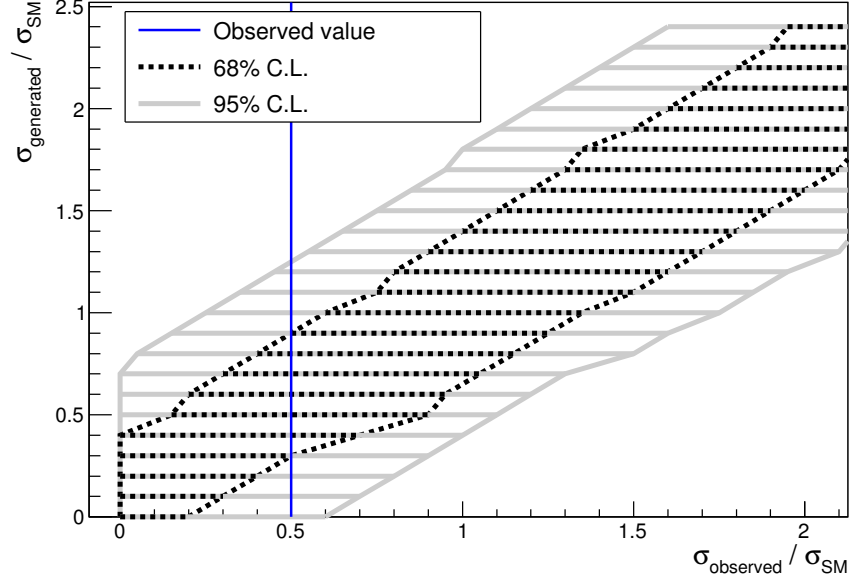


FIG. 10: Confidence bands showing the expected range of observed cross sections as a function of the true cross section, with 68% C.L. (black dashed region) and 95% C.L. (solid gray region).

- [15] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, J. High Energy Phys. 07, 001 (2003).
- [16] T. Sjostrand, S. Mrenna, and P. Z. Skands, J. High Energy Phys. 05, 026 (2006).
- [17] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. D **79**, 052008 (2009).
- [18] A. Bhatti et al., Nucl. Instrum. Methods A **566**, 375 (2006).
- [19] A. Paramonov, F. Canelli, M. D'Onofrio, H. Frisch, and S. Mrenna, Nucl. Instrum. Methods A **622**, 698 (2010).
- [20] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, PoS ACAT, 040 (2007).
- [21] We define the transverse mass as  $m_T = \sqrt{2p_T^{\text{lepton}} \cancel{E}_T (1 - \cos \Delta\phi_{\ell, \cancel{E}_T})}$ , where  $\Delta\phi_{\ell, \cancel{E}_T}$  is the difference in  $\phi$  in the direction of the lepton  $\vec{p}_T$  and  $\vec{\cancel{E}}_T$  vectors.
- [22] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, F. Maltoni, T. Plehn, D. L. Rainwater, and T. Stelzer, J. High Energy Phys. 09, 028 (2007).
- [23] J. Freeman, J. Lewis, W. Ketchum, S. Poprocki, A. Pronko, V. Rusu, and P. Wittich, Nucl. Instrum. Methods A **663**, 37 (2012).
- [24] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. D **82**, 112005 (2010).
- [25] U. Langenfeld, S. Moch, and P. Uwer, Phys. Rev. D **80**, 054009 (2009).
- [26] G. J. Feldman and R. D. Cousins, Phys. Rev. D **57**, 3873 (1998).